



UNITED STATES DEPARTMENT OF COMMERCE  
U.S. Census Bureau  
Washington, DC 20233-0001

## 2020 CENSUS PROGRAM MEMORANDUM SERIES: 2022.06

**Date:** February 25, 2022

**MEMORANDUM FOR:** The Record

**From:** Albert E. Fontenot, Jr. (**Signed February 25, 2022**)  
Associate Director, Decennial Census Programs

**Subject:** The Design of the Post-Enumeration Survey for the 2020 Census

**Contact:** Patrick Cantwell  
Decennial Statistical Studies Division  
301-763-4982  
Patrick.j.cantwell@census.gov

This memorandum releases to the public a planning document, originally finalized in 2019 as an internal document. It provides the design of the Post-Enumeration Survey (PES) to measure the coverage of the 2020 Census. We are publishing this now to provide background for the upcoming release of the first results from the PES.

### The 2020 Census Memorandum Series

The 2020 Census Memorandum Series documents significant decisions, actions, and accomplishments of the 2020 Census Program for the purpose of informing stakeholders, coordinating interdivisional efforts, and documenting important historical changes.

A memorandum generally will be added to this series for any decision or documentation that meets the following criteria:

1. A major program-level decision that will affect the overall design or have significant effect on 2020 Census operations or systems.
2. A major policy decision or change that will affect the overall design or significantly impact 2020 Census operations or systems.
3. A report that documents the research and testing for 2020 Census operations or systems.

Visit 2020 Census on [Census.gov](https://www.census.gov) to access the Memorandum Series, the 2020 Census Operational Plan, and other information about preparations for the 2020 Census.

## The Design of the Post-Enumeration Survey for the 2020 Census

<b>1.0</b>	<b>INTRODUCTION.....</b>	<b>1</b>
1.1	Overview of the 2020 PES Design .....	2
1.2	High-Level Schedule of 2020 PES Activities.....	8
1.3	PES Activities and Census Activities .....	8
<b>2.0</b>	<b>IDENTIFYING PES SAMPLE AREAS.....</b>	<b>9</b>
2.1	BCU Selection Process .....	9
2.3	Small BCU Subsampling.....	11
<b>3.0</b>	<b>INDEPENDENT LISTING.....</b>	<b>12</b>
<b>4.0</b>	<b>INITIAL HU MATCHING AND FOLLOWUP.....</b>	<b>13</b>
4.1	Initial HU Computer Matching.....	13
4.2	Initial HU Before Followup Clerical Matching.....	14
4.3	Initial HU Followup.....	15
4.4	Initial HU After Followup Clerical Matching .....	16
<b>5.0</b>	<b>IDENTIFYING THE SAMPLE HUS.....</b>	<b>17</b>
5.1	Identifying the P-sample HUs.....	18
5.2	Identifying the PI Sample of HUs.....	18
5.3	Identifying the E-sample HUs .....	19
<b>6.0</b>	<b>PERSON-BASED OPERATIONS .....</b>	<b>20</b>
6.1	Person Interview .....	20
6.2	Automated Geocoding .....	22
6.3	Automated Residence Status Coding.....	23
6.4	Person Matching .....	24
6.5	Person Computer Matching .....	25
6.6	Person Before Followup Clerical Operations .....	26
6.6.1	Clerical Geocoding .....	27
6.6.2	Clerical Residence Status Coding.....	27

6.6.3	Person Before Followup Clerical Matching .....	28
6.7	Person Followup .....	28
6.8	Person After Followup Clerical Matching .....	30
<b>7.0</b>	<b>FHU OPERATIONS.....</b>	<b>31</b>
7.1	FHU Computer Matching and Processing .....	32
7.2	FHU Before Followup Clerical Matching .....	33
7.3	FHU Followup .....	34
7.4	FHU After Followup Clerical Matching.....	35
<b>8.0</b>	<b>ESTIMATION.....</b>	<b>36</b>
8.1	Missing Data for P-Sample People.....	36
8.1.1	Imputing Missing Demographic Characteristics for People.....	37
8.1.2	Person-Level Noninterview Adjustment for PES Household Noninterviews .....	37
8.1.3	Imputing the Probability of P-Sample Inclusion for People.....	37
8.1.4	Imputing the Probability of Match Status for People .....	38
8.2	Missing Data for E-Sample People.....	39
8.2.1	Missing Demographic Characteristics for E-Sample People.....	39
8.2.2	Imputing Missing Enumeration Status for People in the E Sample (Net Error).....	39
8.2.3	Imputing Missing Enumeration Status for People in the E Sample (Components of Coverage) 40	
8.3	Person Net Coverage Error .....	40
8.4	Person Components of Census Coverage .....	41
8.4.1	Calibration Adjustment for E-Sample People .....	41
8.4.2	Erroneous and Correct Enumerations for People .....	41
8.4.3	Person Omissions.....	42
8.5	Missing Data for P-Sample HUs.....	42
8.5.1	Imputing Missing Characteristics for P-Sample HUs.....	42
8.5.2	Imputing Missing HU Status for P-Sample HUs.....	42
8.5.3	Imputing Missing Match Status for P-Sample HUs .....	43
8.6	Missing Data for E-Sample HUs .....	43

8.6.1	Missing Characteristics for E-Sample HUs .....	43
8.6.2	Imputing Missing Enumeration Status for E-Sample HUs (Net Error).....	43
8.6.3	Imputing Missing Enumeration Status for E-Sample HUs (Components).....	43
8.7	HU Net Coverage Error .....	44
8.8	HU Components of Census Coverage .....	44
8.9	Uncertainty Estimation .....	45
<b>9.</b>	<b>REFERENCES.....</b>	<b>46</b>

#### **Attachments:**

Attachment 1: High-Level Business Process Model of Key PES Operations  
Attachment 2: Business Process Model for PES Estimation – People in HUs  
Attachment 3: Business Process Model for PES Estimation – HUs



## **The Design of the Post-Enumeration Survey for the 2020 Census**

### **1.0 Introduction**

The Coverage Measurement (CM) Program activities planned for the 2020 Census will provide estimates of net coverage error and components of census coverage for housing units (HUs) and people in HUs. The CM Program goals will be met by conducting a post-enumeration survey, designed to

- Provide measures of net coverage error.
- Produce measures of components of census coverage, including correct enumerations, erroneous enumerations, imputations, and omissions.
- Produce these measures of coverage for demographic groups and key census operations.

To achieve these goals, the 2020 CM Program will design and conduct the 2020 Post-Enumeration Survey (PES). Similar to the 2010 Census Coverage Measurement (CCM) Program and the 2000 Accuracy and Coverage Evaluation (A.C.E.) Program, group quarters facilities and people residing in those facilities are not within the scope of the 2020 CM Program.

The 2020 PES estimation process will continue to use the dual system methodology for net coverage error estimation. This methodology was used in the coverage measurement programs for the 1980, 1990, 2000, and the 2010 censuses (Fay et al. 1988, Hogan 1993, U.S. Census Bureau 2004). As in 2010, logistic regression rather than the traditional post-stratification methodology will be used to produce dual system estimates. In addition, a correlation bias adjustment will be applied to improve the estimates of net coverage error for certain population subgroups.

We will again be using a straightforward estimator for correct and erroneous enumerations—a basic summation of ratio-adjusted survey weights. The ratio adjustments serve to reduce the variance of the estimates. Estimates of omissions will be obtained by subtracting the estimated correct enumerations in the 2020 Census from the dual system estimate.

The purpose of this document is to describe the design of the PES. Subsequent sections of the document present in greater detail the features of each activity. Section 2 provides an overview of the PES sample. Section 3 describes how the PES sample areas are identified. Section 4 details the Independent Listing (IL) operation. Section 5 details the Initial HU Matching and Followup operations. Section 6 describes how the sample HUs are identified. Section 7 details the person-based operations, while Section 8 covers the Final HU operations. Section 9 covers estimation. An overview of the design of the PES is presented in Section 1.1.

There are three attachments. Attachment 1 is a high-level workflow of key PES operations. Attachments 2 and 3 show business process models for PES Estimation operations for people in HUs, and HUs, respectively.

## 1.1 Overview of the 2020 PES Design

This section presents an overview of the key PES activities shown in Attachment 1.

### *Survey Design*

During the survey design phase, key features of the PES are determined, high-level requirements are written, business process models are created, and detailed operational plans are written to provide the framework for the PES operations.

The high-level design of the PES will be similar to the design of the 2010 CCM. As was done in 2010, the PES will make synthetic estimates of net coverage using dual system estimation and direct estimates of correct enumerations and erroneous enumerations.

The Population sample (P sample) and Enumeration sample (E sample) are the two samples used for dual system estimation (DSE). The E sample and the dual system estimates are used to derive the components of census coverage. The P sample is derived from an independent listing of HUs (completely separate from the census) conducted in a probability sample of basic collection units (BCUs) in each state, the District of Columbia (DC), and Puerto Rico. As in the 2000 A.C.E. and the 2010 CCM, any area designated as a Remote Alaska enumeration area will be out of scope for the 2020 PES. The primary sampling unit, as well as the basic work unit for the PES, is the BCU. In 2010, the primary sampling unit and basic work unit was a block cluster. The source of the E sample is the census HUs and census person enumerations in HUs geocoded to the sample of BCUs selected for the P sample.

The PES design in Puerto Rico is the same as the PES design in the U.S. All automated instruments paper forms, and paper materials will be translated in Spanish for use in Puerto Rico. Paper and automated instruments for Independent Listing (IL), Person Interview (PI), and Person Followup will be available in English and Spanish for the 50 States, DC, and Puerto Rico. Paper forms for the Initial HU Followup and the Final HU Followup will only be available in English for the 50 states and DC, and only available in Spanish in Puerto Rico. The instruments and forms developed for use in Puerto Rico will account for the special features of addresses there.

### *Sample Design*

The PES sample design is a general-purpose sample to support the various objectives of the CM Program, which includes the objective of estimating omissions and erroneous enumerations in addition to net error for the 2020 Census. The sample design for 2020 is essentially the same as that used for the 2010 CCM, except there are no plans for the sample reduction of medium and large BCUs.<sup>1</sup> Additional changes are that the initial primary sampling units are BCUs instead of block clusters and the thresholds for the large stratum changed. For a more thorough discussion

---

<sup>1</sup> An initiative to reduce nonsampling error in the 2010 CCM program included reducing the sample size after the CCM Independent Listing (Whitford 2009). Entire block clusters were systematically removed from the listing sample.



of the 2010 CCM sample design, see Konicki (2010). For information on the 2020 PES sample design, see Hill, Nguyen, and Davis (2018).

The PES sample design includes a number of distinct processes, from creating the sampling frame and selecting sample of BCUs, to selecting addresses for the P and E samples. After the PES BCUs are selected, an address list is created independent of the census for each BCU (this is the IL operation described later).

### *Initial BCU Sample*

The first production operation is the sampling of BCUs. The frame contains all BCUs that are outside of Remote Alaska and not fully covered by water. This list of BCUs is stratified by state, size, percent of homeowners, and an American Indian Reservation indicator. A stratified systematic sample of about 10,100 BCUs is selected in the U.S. An additional 400 BCUs will be selected in Puerto Rico.

The people and HUs in the PES sample of BCUs are independently enumerated twice, once for the P sample and once for the E sample. Both the P sample and the E sample measure the same HU and household population. However, the P-sample operations are conducted independently of the census. After matching with the census lists and reconciliation, the P sample provides information about the population missed in the census, whereas the E sample provides information about erroneous census inclusions. This information is used in different ways to estimate the net coverage error and the components of census coverage.

### *Independent Listing*

In the BCUs selected in the first phase of sampling, the IL is conducted by personal visit using the Listing and Mapping Instrument (LiMA). This operation lists all HUs or potential HUs in each PES sample BCU. Information regarding the number of HUs belonging to a particular structure is also collected. Overall, approximately 564,500 HUs are expected to be listed (540,000 HUs in the U.S. and 24,000 HUs in Puerto Rico).

### *Small BCU Subsample*

PES IL results and preliminary census HU counts are used to stratify and determine the subsampling rates for small BCUs. Small BCUs are subsampled using variable rates to minimize variance on the coverage estimates and increase field efficiency.

### *Initial Housing Unit Matching and Followup*

The Initial HU (IHU) Computer Matching follows, in which the listing results from the PES IL are matched against a preliminary listing of census HU and group quarters addresses within each

sample BCU and one ring of surrounding BCUs.<sup>2</sup> Duplicate addresses<sup>3</sup> within the PES IL and within the preliminary census address listing are also identified.

The next operation is the Initial Housing Unit (IHU) Before Followup (BFU) Clerical Matching, which attempts to resolve possible matches, nonmatches, and duplicates from the IHU Computer Matching. For this operation, the clerical matchers use the PES Census Matching and Map Update (CMMU) system, a computer application that facilitates and records the clerical review and coding of address records. Clerical matchers also search for duplicate addresses within the PES IL and preliminary census address lists, in addition to those duplicates identified during IHU Computer Matching. None of the results from any of the IHU operations are used to update any census information.

Once these initial matching activities are completed, the IHU Followup (IHUFU) field operation attempts to resolve any differences between the PES IL results and the census preliminary address file, and also to resolve potential duplicates identified in the computer or clerical matching. An attempt is also made to resolve HU status where IL indicated the unit had a status other than occupied or vacant at the time of listing. For example, the unit may have been under construction. The IHUFU operation will be conducted using paper forms.

When the IHUFU is completed or not needed, the IHU After Followup (AFU) Clerical Matching ensues for the BCU. In the IHU AFU, results from the IHUFU are used to assign the match status, duplicate status, or HU status for unresolved addresses. This process produces files that contain match codes for PES and census addresses in the sample BCUs and in the surrounding BCUs.

### *Person Interview Sample Identification*

The next PES operation selects a sample of addresses in large BCUs. This sampling reduces the cost of the PES and reduces the clustering of the PES sample, so that a greater number of BCUs can be included in sample. As BCUs complete the IHU matching, a Preliminary Enhanced List (PEL) of addresses is sent to the large BCU subsampling operation for Person Interview (PI). The addresses on the PEL will be sampled in five waves so that the PI doesn't have to wait until all IHU matching operations end. Note that some BCUs will skip IHUFU and IHU AFU, so some PELs will be created before IHU BFU finishes. The PEL provides the frame from which the sample is selected. Only large BCUs, those with 58 HUs or more, are eligible for subsampling. Once HUs in the large BCUs are sampled out, the remaining sample cases are sent to the PI. The PI workload has two components: those addresses eligible for and selected to be in the P sample and those addresses not eligible to be in the P sample. These latter addresses are also referred to as "census-only" HUs.<sup>4</sup>

<sup>2</sup> A surrounding BCU is one that touches the BCU at one or more points. This definition includes the BCUs that touch the corner of the BCU.

<sup>3</sup> Duplicate addresses refer to duplicate address records that reference the same living quarters within a given address list.

<sup>4</sup> Units eligible to be in the P sample must have been listed during the PES IL operation. Census-only units are census units geocoded to the sample BCU that are determined to be valid HUs missed in the PES IL or actually



After selecting the PES P-sample addresses for interviewing, the 2020 PES sample is expected to be approximately 180,000 HUs for the U.S. and 8,000 HUs for Puerto Rico (see Table 2). The national sample is distributed among the 50 states and D.C. roughly proportional to population size, although there are slight increases in the sample for small states, Hawaii, and American Indian Reservations.

Table 2 shows the expected size of the 2020 PES sample in terms of the expected number of HUs and sample BCUs.

**Table 2: 2020 PES Expected Sample Sizes**

Geography	Independent Listing Sample		P-Sample HUs
	BCUs	HUs	
U. S.	10,100	540,000	180,000
Puerto Rico	400	24,000	8,000

### *Person Interview*

For the selected HUs in each sample BCU, interviewers employ a computer-assisted data collection instrument on a laptop to obtain information about the residents of the sample HU on Interview Day (the day the PI is conducted) and people who moved out of the sample HU between Census Day (April 1, 2020) and Interview Day (called outmovers).

Including the census-only HUs and outmovers in the PI reduces the followup workload (since they could eventually be selected for the E sample). Including the outmovers also facilitates matching to the census enumerations for estimating coverage components.

The 2020 PES is using Procedure B, whereby Interview Day residents of the sample HU are included in the P sample. The 2010 CCM employed Procedure B+, whereby Interview Day residents of the sample HU and certain people who moved out of the sample HU between Census Day and Interview Day were included in the P sample. The 2000 A.C.E. employed the PES C methodology where the P sample included Census Day residents of the sample HU (nonmovers and outmovers), as well as people who moved into the sample HU after Census Day (inmovers). For 2000 A.C.E. only the Census Day residents of the HU were eligible for matching, however.

### *E-sample Identification*

A primary goal of E-sample identification is to identify changes (primarily adds and deletes) to the census HUs in sample BCUs between address canvassing and the final census universe. The E sample is the set of person enumerations and HU enumerations in the census within PES sample BCUs. The E sample is necessary to estimate the correct and erroneous enumerations in the census. Census adds are HUs that were added to the final census universe but were not included as valid HUs in the initial census universe, before mailing paper census forms. Census

---

located within the surrounding BCUs to the sample BCUs. Census-only units are not eligible for the P sample but are likely to be in the E sample.

adds are identified so they can be given a chance to match to P sample HUs and other census HUs. In the E-sample identification, census units that fall in PI sample segments are identified in a way to assure that the E sample and P sample are in common areas. This improves the precision of the dual system estimates and the cost-effectiveness of succeeding field operations. It also reduces the number of E-sample cases requiring followup.

### *Person-based Operations*

Before the Person Matching operations begin, an automated operation to assign a residence status code to all people listed in the PI is conducted. Also, an automated operation to assign geocodes to alternate and in-mover addresses<sup>5</sup> collected in the PI is conducted. This is followed by clerical operations conducted to 1) assign geocodes to respondent-provided addresses and 2) assign residence status codes to people when these codes could not be assigned during the automated coding operations.

For 2020, the PES Person Computer Matching will attempt to search for matches between people rostered at the sample addresses during the PI and people enumerated in the census in the PES sample BCUs and surrounding BCUs. Alternate and in-mover addresses collected during the PI and geocoded during automated or clerical geocoding are also used to identify other places to search for matches between PI and census enumerations. The computer matching will also conduct a nationwide search for matches between the PI and census enumerations. Matching people between Puerto Rico and the 50 states and DC is out of scope. A computerized search for duplicate enumerations in either the PI or census is also conducted. For PI, the search is within each sample BCU. For duplicate person listings enumerated in the census, we search both within the sample BCUs and either nationwide for U.S. residents or within Puerto Rico for Puerto Rico residents.

The Person BFU Clerical Matching follows in which the matching staff assign the status of match, possible match, or nonmatch to PI and census person records. Clerical matchers will use a computer application, developed specifically for PES clerical matching, that facilitates and records the clerical review, matching and coding of the person data. Other information used in this activity includes the PES and census maps as well as the results of the previous computer matching. A clerical search for duplicates is also performed. Cases remaining unresolved in terms of match status, enumeration status, or residence status become eligible for the next PES operation, Person Followup.

During PES Person Followup (PFU), interviewers in the field contact cases within the PES sample BCUs identified in the Person Matching operations with unresolved residence or enumeration status, even when they may have been matched, to resolve the particular issue. Personal visits will be conducted to follow up on any possible matches or possible duplicates where one of the records in the linked pair is located outside the PES sample BCU or ring of

---

<sup>5</sup> Respondent-provided addresses are any addresses, other than the PES sample address, where someone lived or could have been counted by the census. There are two types of respondent-provided addresses, in-mover addresses and alternate addresses. The Census Day residence for in-movers is referred to as the in-mover address. Other addresses provided by the respondent, whether an in-mover or non-mover, are classified as alternative addresses.



surrounding BCUs. Thus, the PFU is sometimes conducted outside the sample BCU to follow up on links found outside the BCU during the nationwide search. The PFU interview is completed using a paper questionnaire.

After PFU is complete, the Person AFU Clerical Matching uses the information obtained from that interview to attempt to resolve match, residence, or enumeration status for people sent to PFU.

### *Final Housing Unit Matching and Followup*

Once all of the person-based operations have been performed, final HU activities are conducted to determine the final match codes assigned to P-sample and E-sample addresses. These results will be used to estimate census coverage of HUs. The first of these tasks is the Final HU (FHU) Computer Processing. In this operation, a determination is made as to which HUs will proceed to the FHU BFU Clerical Matching operation, based on changes to the lists of PES or census units since the IHU Matching and IHUFU operations were conducted.

During FHU BFU Clerical Matching, the staff use computer-assisted clerical matching techniques, as well as PES and census maps, to review and assign the status of match, possible match, or nonmatch to addresses sent from the FHU Computer Processing. During FHU Clerical Matching, the matchers attempt to determine whether or not an address should have been counted as a valid HU on Census Day. The FHU Clerical Matching operation is similar to the Initial HU Clerical Matching operation described previously, except that it uses FHU census data rather than data from the preliminary list of census addresses. Also, only those PES IL addresses that were selected for PI are included in the FHU operations. Results from the Person Matching and PFU operations are available to assist the matching staff, as well as results and data from the IHU matching and followup operations.

Those cases that need additional information are candidates for the FHU Followup (FHUFU) operation. In this operation, attempts are made to collect information to resolve any remaining differences between the P- and E-sample HUs. The FHUFU data collection forms are paper-based.

The concluding step in the FHU matching operations is the PES FHU AFU Clerical Matching. When conducting this operation, results from the FHUFU (and any previous PES Person and HU Matching operations) are used to resolve any outstanding issues regarding match, HU, or duplication status. Once FHU AFU has concluded, the PES HU Estimation process begins.

Each of the field and clerical matching operations includes a quality control procedure. The quality control check for the field operations consists of field supervisors observing interviews, reviewing and editing data, and reinterviewing selected HUs. The clerical matching quality operations include reviews by technicians of the computer processing, computer matching, and followup results and reviews by analysts of the work done by the technicians. More information on the quality control program for each operation is covered in later sections.



## Estimation

The PES Estimation process consists of several operations, which will ultimately lead to the production of estimates of census coverage for both HUs and people in HUs. This includes estimates of net coverage error and estimates of the components of census coverage, including correct enumerations, erroneous enumerations, whole-person census imputations, and omissions. As part of this estimation, we will implement procedures to account for missing data and reduce the sampling and nonsampling errors in the coverage estimates.

### 1.2 High-Level Schedule of 2020 PES Activities

A high-level schedule of the 2020 PES activities described in the previous section is as follows:

**Table 1: High-Level Schedule of 2020 PES Activities**

Activity <sup>6</sup>	Planned Start Date	Target End Date
Select BCU Sample	July 28, 2019	August 16, 2019
Conduct Independent Listing	January 16, 2020	March 14, 2020
Conduct Small BCU Subsampling	March 31, 2020	April 17, 2020
Conduct Initial HU Computer Matching	April 1, 2020	April 24, 2020
Conduct Initial HU Before Followup Clerical Matching	April 15, 2020	May 15, 2020
Conduct Initial HU Followup	May 6, 2020	June 13, 2020
Conduct Initial HU After Followup Clerical Matching	June 2, 2020	July 24, 2020
Identify Person Interview Sample <sup>7</sup>	May 15, 2020	August 14, 2020
Conduct Person Interview	June 17, 2020	September 12, 2020
Conduct Clerical Geocoding, and Residence Status Coding	August 3, 2020	November 20, 2020
Conduct PES PI Duplicate Search	November 9, 2020	November 20, 2020
Identify E Sample	November 30, 2020	December 22, 2020
Conduct Person Computer Matching	December 7, 2020	December 30, 2020
Conduct Person Before Followup Clerical Matching	January 7, 2021	February 19, 2021
Conduct Person Followup	February 3, 2021	March 20, 2021
Conduct Person After Followup Clerical Matching	February 25, 2021	April 16, 2021
Conduct FHU Computer Matching	March 15, 2021	April 20, 2021
Conduct FHU Before Followup Clerical Matching	April 26, 2021	May 21, 2021
Conduct FHU Followup	May 19, 2021	June 12, 2021
Conduct FHU After Followup Clerical Matching	June 7, 2021	July 23, 2021
Produce PES National Person Demographic Estimates	April 16, 2021	June 30, 2021
Produce PES State and National Person Operational Estimates	April 30, 2021	September 30, 2021
Produce PES HU Estimates	July 23, 2021	September 30, 2021

### 1.3 PES Activities and Census Activities

Attachment 1 details the relationship between the aforementioned PES activities and 2020 Census activities. PES uses the results of four census processes. First, information from the

<sup>6</sup> This schedule does not include additional time for quality control and reinterview. In general, quality control and reinterview start a week after the planned start date and continue for a week after the target end date.

<sup>7</sup> Six separate samples will be selected for the Person Interview during this period.

Master Address File is used to construct the PES BCU sampling frame. Second, to conduct the Initial HU Computer Matching and conduct the subsampling for small BCUs, the PES relies on the preliminary census address list from Address Canvassing. Next, to identify the E sample and conduct person and HU matching and PFU, the PES requires the unedited person, HU, and group quarters (GQ) data files obtained from the census (Census Unedited File or CUF). Lastly, the PES needs census edited person and HU data (the Census Edited File or CEF) to do the estimation.

In addition, to ensure independence between census and the PES, the PI data collection is scheduled to start after most of the Nonresponse Followup (NRFU) operation is complete in the BCU.

## **2.0 Identifying PES Sample Areas**

There are two major sampling activities to identify the areas in which PES operations will be conducted. These activities are: selecting the BCUs where the PES IL operation will take place and subsampling the small BCUs.

### **2.1 BCU Selection Process**

Within each state and Puerto Rico, BCUs are stratified based on the cross-classification of size, tenure (owner/non-owner<sup>8</sup>), and American Indian Reservation status. First, BCUs are classified by size into three mutually exclusive groups based on the expected number of HUs within the BCU. The expected number of HUs within the BCU is derived by counting addresses on the Master Address File. The three groups are:

- small BCUs – 0 to 2 HUs
- medium BCUs – 3 to 57 HUs
- large BCUs – 58 or more HUs

The second classification categorizes medium and large BCUs based on tenure, i.e., for the entire BCU, the proportion of households that are either rented or owned based on Census 2010 data. BCUs where more than 40 percent of the households do not own their home are placed into the non-owner stratum.

For 26 states, an American Indian Reservation stratum is defined. This stratum contains BCUs with 3 or more HUs located on an American Indian Reservation or associated trustland, regardless of the tenure status of the BCU. To adequately control sample sizes for American Indians, inclusion in the AIR stratum takes precedence over the other stratification variables. Table 3 summarizes the six mutually exclusive sampling strata.

---

<sup>8</sup> Technically, the non-owner category includes renters and persons occupying HUs without payment of cash rent. However, renters make up the vast majority of people in the non-owner category, so we will refer to this category as “renters” throughout the document, unless we are referring to a specific sampling stratum.

**Table 3: Strata for Selecting the BCU Sample for PES Independent Listing**

First-Phase Stratum	Size Definition	Tenure Definition
(1) Small	0-2 HUs	NA
(2) Medium Owner	3-57 HUs	BCUs with less than 40% of non-owner households
(3) Medium Non-Owner	3-57 HUs	BCUs with 40% or more of non-owner households
(4) Large Owner	58+ HUs	BCUs with less than 40% of non-owner households
(5) Large Non-Owner	58+ HUs	BCUs with 40% or more of non-owner households
(6) AIR (26 states only) <sup>1</sup>	3+ HUs	NA

<sup>1</sup> These are the states with sufficient American Indian population living on AIR in 2010.

Note: There may be AIR in other states, but with insufficient American Indian population to support a separate AIR stratum.

NA Not applicable.

AIR is American Indian Reservation.

**Allocating the Sample.** The IL BCU sample in the U.S. is a national sample of about 10,100 BCUs: 8,000 medium and large BCUs and 2,100 small BCUs. Included in the 8,000 BCUs is a separate sample of BCUs for American Indian Reservations. The number of medium and large BCUs is based on an expected average of P-sample HUs per BCU in each state from preliminary housing unit counts by BCU for each state.

The 8,000 medium and large BCUs are allocated to the 50 states and DC. This allocation is proportional to the population in each state with three exceptions. First, Hawaii is allocated enough BCUs to support an expected HU sample size of 4,500 to help the reliability of the Native Hawaiian or Pacific Islander estimates. Second, each state is allocated a minimum sample size to get an expected 1,050 HUs, not including housing units in the small and American Indian Reservation strata. Third, the states with American Indian Reservation stratum are allocated enough BCUs to support an expected 10,500 HUs in the American Indian Reservation stratum. The allocation of BCUs to the 26 American Indian Reservation states is proportional to the 2010 Census population of American Indians living on reservations on each state. The allocation of the 2,100 small BCUs to the states is proportional to the number of small BCUs in each state.<sup>9</sup>

For Puerto Rico, the expected number of sample BCUs for IL is 400: 280 medium and large BCUs and 120 small BCUs.

**Probabilities of Selection.** To help support the reliability of minority and renter estimates, the sampling operation oversamples BCUs with higher proportions of renter households. This oversampling will help increase the sample size of minorities and renters. Also, the large BCUs are selected with higher probability than medium BCUs to allow for the subsampling of HUs within these BCUs later on. This brings the overall probability of selection for HUs in large BCUs more in line with the overall probability of selection for HUs in medium BCUs. Oversampling of large BCUs increases the number of unique BCUs in sample to provide a more geographically diverse HU sample.

<sup>9</sup> See Hill, Nguyen, and Davis (2018) for sample allocation details.



## 2.3 Small BCU Subsampling

Using results from the IL as well as preliminary census address data<sup>10</sup>, BCUs from the original small sampling stratum (0-2 HUs) are stratified into four groups based on the updated number of HUs per BCU. The updated HU count for each BCU is the larger of the number of HUs from the PES IL operation and the number of HUs from the preliminary census address list. The subsampling is designed to minimize the impact on the variance of the coverage estimates and to increase the efficiency of the PES field operations.

The strata are formed so that small BCUs with more than the expected number of HUs will be sampled at a higher rate to keep their weights lower. This will reduce the potential contribution from these BCUs to the sampling variance of the coverage estimates.

Table 4 displays the four strata used for small BCU subsampling, as well as their target percent retained. Approximately 10 percent of the original small BCUs with 0 to 2 HUs found is retained in the final BCU sample and 30 percent of the small BCUs with 3 to 9 HUs found is retained in sample. The target national BCU sample size is allocated to the 50 states and DC proportional to the HUs in each state.

**Table 4: 2020 PES Small BCU Subsampling Strata**

Second-Phase Stratum (BCU size)	2020 Target Percent Retained
0 to 2 HUs	10
3 to 9 HUs	30
10+ HUs	100
AIR*/Trustland	
TJSA/TDSA/ANVSA**	100

\*AIR is American Indian Reservation.

\*\* Tribal Jurisdiction Statistical Area/Tribal Designated Statistical Area/Alaska Native Village Statistical Area.

As shown in Table 4, small BCUs with 10 or more HUs are not subsampled. Also, BCUs which are part of an American Indian Reservation, associated trustland, TJSA, TDSA, or ANVSA (see Table 4 caption) are not subsampled. Within each of the other two strata, BCUs are sorted and a systematic sample of BCUs are selected with equal probability within each subsampling stratum within each state, DC, and Puerto Rico.

<sup>10</sup> Same preliminary census address data used in the Initial HU Matching discussed in Section 5.1.

### 3.0 Independent Listing

An independent listing of all HUs and potential HUs in the PES sample BCUs is compiled before Census Day. The listers make a complete canvass of their assigned BCUs to collect the following information for each HU and potential HU:

- address (information to identify the HU and structure, street name, city, and zip code) for units with city-style<sup>11</sup> addresses and for non-city-style when available.
- rural route and box number (for both non-city-style and city-style addresses).
- respondent name and physical location description for units with non-city-style addresses.
- number of HUs in any multi-unit structure.
- description of each building that contains HUs.<sup>12</sup>
- status of each unit listed such as occupied, vacant and intended for occupancy, under construction, future construction, unfit for habitation, boarded up, empty trailer lot or site, structure used for storage of household goods, or structure used for another purpose.
- special features of addresses in Puerto Rico will also be captured.

Independent listers will use the Listing and Mapping Instrument (LiMA) to conduct the operation.

To get information about extra and hidden units, listers will attempt to contact a respondent at each address. Listers will make two attempts at the unit before contacting a proxy and use observation only when necessary. In multi-unit structures, the lister will canvass by observation and then attempt to gather and reconcile the information from a manager who may be able to provide information on all HUs more efficiently than questioning residents of each unit directly. Also, efforts will be made to acquire maps of apartment complexes, mobile home parks, and recreational vehicle (RV) parks from the managers of these facilities.

Units in RV parks, campgrounds, marinas, and other similar locations will be listed only if someone with no other usual residence occupies them at the time of listing. In addition, vacant lots in mobile home parks will be listed since they could contain a HU when the PES Person Interview is conducted. If the listers are unsure about whether or not a unit is a HU, they will be instructed to list the addresses. The assignment of listers to BCUs adheres to strict rules for independence between census and PES operations (Cantwell, 2018).

Listers will use the LiMA to place a map spot on the map to indicate the location of each structure listed.

The quality assurance plan for Independent Listing has four components:

- All listers undergo online and classroom training.

---

<sup>11</sup> City-style is a form of address where house number and street name are given (e.g., 123 Blue Jay Road). Non-city-style addresses pertain to all addresses other than city-style.

<sup>12</sup> Description of a building includes identifying the building as a single-family house; building with two or more apartments; mobile home or trailer (inside or outside of a mobile home park); camper, tent, boat, van; or other.



- Listers who struggled in training are observed by their field supervisor.
- Automated instrument edits to detect and correct common data-entry errors. Some edits prevent the lister from moving forward until a correction is made, while other edits warn the lister of a possible error but allow them to move on without a correction.
- Some BCUs are sent for in-field Quality Control (QC), as described below.

When a BCU is completed during production, it is run through a series of four checks to place the BCU into one of three sampling strata. All BCUs in the high sampling stratum are selected for in-field QC. Ten percent of BCUs in the middle sampling stratum and five percent of BCUs in the low sampling stratum are systematically sampled for in-field QC. Once a BCU is selected for in-field QC, another lister specifically hired for QC will check a sample of addresses in that BCU to verify addresses or make corrections, as needed. After the sample is complete, the automated instrument calculates a pass or fail decision. If the production lister made fewer than the number of allowable errors, the BCU passes QC. Otherwise, the BCU fails QC and the QC lister reworks all remaining addresses in the BCU.

#### **4.0 Initial HU Matching and Followup**

The PES Initial HU operations contain several activities, all of which are designed to assign HU status and match status. The Initial HU Computer Matching and Initial HU Before Followup (BFU) Clerical Matching attempt to match units from the IL to preliminary census addresses. The results of these two matching processes are recorded so that units from the IL and census units that remain unresolved can be sent to the next operation, Initial HU Followup (IHUFU). In IHUFU, information is collected that may explain any differences between the IL results and the preliminary census addresses, while also attempting to resolve potential duplicates. This information is used during the Initial HU After Followup (AFU) Clerical Matching to assign final match codes from the Initial HU Matching and Followup operation.

##### **4.1 Initial HU Computer Matching**

During Initial HU Computer Matching, all addresses listed in the PES sample BCUs during IL are computer matched to preliminary census addresses (of HUs and GQs) within the sample BCUs and one ring of surrounding BCUs. This matching operation will identify:

- matches and possible matches between the IL addresses and census addresses.
- potential PES IL duplicates within the sample BCU.
- potential census duplicates.

The data from the LiMA and the preliminary census addresses go through a series of data preparation steps before matching, including address standardization. Addresses from either file that are blank or could not be standardized are not subject to computer matching, but will be included in the clerical matching operation.

Potential duplicates are identified through a one-to-many matching process. Links between a single census address and multiple IL addresses identify a potential PES duplicate; links between a single PES IL address and multiple census addresses identify a potential census duplicate.



The results of the computer matching and additional information are loaded into a HU clerical matching system, which is subsequently used by clerical matchers. All of the results from computer matching will be checked clerically.

## **4.2 Initial HU Before Followup Clerical Matching**

The Initial HU BFU Clerical Matching procedure uses the results of the computer matching stage in an attempt to further match and review HUs from both the PES IL and the census preliminary address file. There are two stages of Initial HU BFU Clerical Matching: BFU Technician and BFU Analyst.

In the BFU Technician stage, technicians use clerical matching software to:

- review the results from the Initial HU Computer Matching operation to determine the correct match status and to unlink addresses as necessary.
- attempt to match IL units and census units that were not matched in the Initial HU Computer Matching operation.
- search for duplicates:
  - between IL units within the PES sample BCUs.
  - between units on the preliminary census list in the PES sample BCU and the surrounding BCU.
- assign match codes to indicate the results of their review.

The supplemental materials used in the clerical matching include electronic maps with the PES map-spotted units and electronic census maps. Given the number of non-city-style addresses in rural areas, map spots and descriptions are very important for the clerical matching procedure in these areas.

In the IHU BFU Analyst stage, analysts perform the QC of the technicians' work and review cases referred by the technician. The referred cases are IL or census units where the match or duplicate status is not clear. The QC process for clerical matching consists of analysts performing a full review of BCUs selected for the matching QC sample. For the IHU BFU Clerical Matching operation, approximately 33 percent of all BCUs going through the operation will be checked. The target Average Outgoing Quality Limit (AOQL) is 3.5 percent.

After training, technicians will start production work in 100% QC status. If a technician is performing poorly, the technician may be required to undergo retraining. Technicians that perform poorly even after retraining will be removed from the operation.

To help ensure the quality of the matching and ensure the newer analysts understand procedures, the experienced analysts will informally review the work of the newer analysts.

Some BCUs meeting specific criteria will be allowed to skip the BFU Clerical Matching and go directly to IHUFU. Other BCUs will be allowed to skip both BFU Clerical Matching and the

IHUFU.<sup>13</sup> All BCUs skipping BFU Clerical Matching will be subject to a clerical duplicate search in Initial HU AFU Clerical Matching.

The product of the BFU Clerical Matching process will be a file containing the Initial HU match codes for all addresses (both IL and census), and flags indicating which census or IL units will go to the next operation, IHUFU.

### 4.3 Initial HU Followup

For the IHUFU operation, units from either the IL or census that meet pre-determined criteria are followed up in the field to determine HU status or match status. The following are the conditions that will render a unit eligible for followup:

- census unit not linked to an IL unit in the same PES sample BCU.
- IL unit not linked to a census unit in the PES sample BCU or surrounding BCUs.
- IL or census unit with a link assigned a possible match status.
- IL unit with unresolved HU status (under construction, future construction, unfit for habitation, vacant trailer pad, or other), regardless of match status.
- IL unit with a duplicate link, regardless of match status.
- IL unit with a link to a census unit in a surrounding BCU, regardless of match status.
- census unit with a duplicate link, regardless of match status.
- IL unit with a link to a census GQ.

To maintain the independence between the PES and census, previous census and PES work assignments are taken into account when making assignments to PES staff. As with all other PES followup activities, staff will use a paper form seeking to acquire the necessary information. Also, PES field representatives will have access to the LiMA with both PES and census maps. Further, they will be able to update maps in limited situations.

The followup forms are customized for each individual situation. Efforts will be made to collect the following information:

- BCU number (for a unit that could be a duplicate in a surrounding BCU).
- Evidence to determine if there was a HU at the address on the date of the followup visit.<sup>14</sup> If the case in question is not a HU, a statement should be provided as to the reason for this determination.
- Whether or not two units identified as possible matches are the same unit.
- Whether or not two or more units identified as possible duplicates are the same unit.
- HU status updates for certain situations.

---

<sup>13</sup> One example of the skip criteria: If all IL and all census addresses in a BCU have a match code of “match” and all of the IL units have a unit status code of “occupied or vacant intended for occupancy,” the BCU can skip both the Initial HU BFU Clerical Matching and the IHUFU.

<sup>14</sup> It is possible that some of the followup cases were not HUs at the time of listing. In this situation, the interviewer will need to determine if a HU does actually exist at the address.



- Response to case-specific questions provided by the clerical matchers.

Also, field representatives will use the reference list<sup>15</sup> provided to determine if any units sent for followup match other units on the reference list. This is done to reduce the number of potential E-sample nonmatches. Cases lacking sufficient information to locate the unit in the field and conduct the followup interview will not go to IHUFU.

The IHUFU quality assurance plan has the following components:

- All listers undergo online and classroom training.
- Listers who struggled in training are observed by their field supervisor.
- Field supervisors review and edit completed forms for accuracy and completeness.
- A separate staff of listers, called QC listers, perform a check of data accuracy for randomly selected IHUFU forms in each BCU. Approximately 15 percent of all cases sent to IHUFU will be checked in the field, with a target AOQL of 4.0 percent.
- A field supervisor edits the QC forms for accuracy and completeness.
- Each BCU undergoes an office edit.

#### **4.4 Initial HU After Followup Clerical Matching**

Once the IHUFU is completed, the IHUFU paper forms, along with Initial HU BFU Clerical Matching results file and updated PES maps, are made available for the next PES operation, the Initial HU AFU Clerical Matching. There are three stages in the Initial HU AFU Clerical Matching operation: AFU Technician, AFU Analyst, and AFU Outlier Review.

A main objective of the AFU Technician stage is to have the technicians assign match status and HU status codes to cases that were sent to IHUFU. A clerical search is performed to attempt to find duplicates for BCUs that have skipped Initial HU BFU Clerical Matching.

In addition, clerical matchers will correct addresses and update PES maps by adding or deleting map spots, based on information collected in IHUFU. Clerical matchers may add map spots to the PES maps for census addresses that belong within the sample BCU or a surrounding BCU and were listed in the sample BCU by census, but were not listed there in IL. Clerical matchers may also delete map spots for IL addresses that do not belong in the sample BCU or do not represent a unit that has the potential to be a HU by the time of the PI, such as businesses. These address corrections and map spot numbers are recorded and are used for future PES field and matching operations, but are not fed back into any actual census operation.

The IHU AFU Analyst stage for IHU AFU Clerical Matching is very similar to that for IHU BFU Clerical Matching. In this stage, the analyst conducts a quality check on a sample of technicians' work and reviews any cases referred by the technician. The expected QC workload and target AOQL for this operation are the same as for IHU BFU Clerical Matching.

---

<sup>15</sup> The reference list is a list of all of the PES HUs and census HUs and GQs in the PES sample BCU. It also includes census units in surrounding BCUs that contain a unit linked to a unit in the sample BCU.



After training, technicians will start production work in 100% QC status. If a technician is performing poorly, the technician may be required to undergo retraining. Technicians that perform poorly even after retraining will be removed from the operation.

To help ensure the quality of the matching and ensure the newer analysts understand procedures, the experienced analysts will informally review the work of the newer analysts.

In the AFU Outlier Review stage, analysts review BCUs identified for review by Headquarters staff and BCUs that exceed a pre-specified threshold of weighted nonmatch and erroneous enumeration counts. The analysts review the BCUs for systematic errors or issues. They document their findings in journals that can later be used during estimation processes if questions arise.

The output for the Initial HU AFU Clerical Matching operation consists of the results file, containing the match status codes and HU status codes of all census and IL units included in the initial HU operations.

## **5.0 Identifying the Sample HUs**

Identifying the P- and E-sample HUs are separate sampling activities done at different times. Along with identifying the P-sample HUs, an additional sample of census-only HUs is identified to receive the PI. Since these census-only units were not listed in the PES IL they cannot be part of the P sample, but a sample of them is selected to be included in the PI, along with the P-sample HUs. Identifying the PI sample involves selecting a subsample of HUs within large BCUs.

In large BCUs (58 or more HUs), the HUs within the BCU are subsampled. This feature of the sample design is intended to increase the number of BCUs in sample while still attaining the targeted number of HUs for the P sample. The objective is to reduce costs and yield manageable field workloads for PES PI and PFU without seriously affecting the precision of the PES coverage estimates. Also, by taking advantage of the homogeneity of units within the large BCUs, interviewing a manageable fraction of several different large BCUs provides a more geographically diverse sample. We expect this to have lower sampling variance than a design with more sample housing units within BCUs and fewer BCUs.

When selecting the initial sample of PES BCUs for IL, large BCUs have a higher selection probability than medium BCUs to take into account the anticipated subsequent HU subsampling. After the subsampling of HUs in large BCUs, the HU selection probabilities in medium and large BCUs in the original BCU sampling strata are similar.

Another important goal of the HU subsampling is to geographically overlap the P and E samples to reduce the E-sample PFU workload. An overlapping P and E sample is not necessary but improves the precision of coverage estimates, the cost-effectiveness of the succeeding field operations, and the data processing efficiency.

## 5.1 Identifying the P-sample HUs

The PES independently listed units that are determined to be HUs or potential HUs and which are located in the final PES sample BCUs after the IHU Matching and Followup operations are the source of the P sample. For BCUs containing 57 or fewer IL HUs, all the PES HUs are included in the P sample. For BCUs with 58 or more IL HUs, a subsample of PES HUs is selected. All HUs in BCUs in the American Indian Reservation stratum are included in the P sample.

We subsample HUs within the large BCUs by forming groups of adjacent HUs, called segments, and selecting one or more segments per BCU for the P sample. The segments are constructed by sorting housing units by map spot number within BCU. The map spot numbers are reviewed during the Initial Housing Unit matching operations to assure they are sequential numbered around the BCU. The segments have roughly equal numbers of HUs within a BCU. Segments of HUs are used as the sampling unit to obtain compact interviewing workloads and to facilitate the identification of an overlapping E sample. A systematic sample of segments is selected across the BCUs within a stratum. The resulting sample will yield at least one segment selected per BCU. The PES IL units in BCUs where no subsampling occurs and the PES IL units that are retained after the subsampling compose the P sample. The expected numbers of P-sample HUs are shown in Table 2.

As was done in the 2000 A.C. E., the PI sample will be selected on a flow basis for the 2020 PES. As BCUs complete Initial HU Clerical Matching, the addresses that are eligible for the PI shall be input to PI sample selection. Overall, we expect to sample addresses for PI in five waves spread across the P-sample selection period.

## 5.2 Identifying the PI Sample of HUs

The PI sample is the set of HUs where the PES PI is conducted. It consists of the HUs selected for the P sample (described in the previous section) and a sample of census-only units. Census-only HUs are eligible for the PES PI to obtain Census Day residence status of people in these HUs prior to identifying E-sample HUs. There are four categories of census-only HUs eligible for the PI:

- Nonmatched census unit that is confirmed to exist in the sample BCU, but was missed by the PES IL.
- Nonmatched census unit that is geocoded to the PES sample BCU, but confirmed to exist in one of the BCUs surrounding the sample BCU.
- Nonmatched census unit geocoded to the PES sample BCU, with an unresolved HU status code.
- Matched census unit that is geocoded to the sample BCU, but confirmed to exist in one of the BCUs surrounding the sample BCU. (The corresponding IL unit is not eligible for the P sample, since it is not located in the sample BCU.)

Census-only HUs are eligible for the E sample if they are on the final census address list for the final sample BCU.



For PI sample identification, the eligible census-only HUs are assigned to the same segments described in Section 6.1 by sorting all PES IL HUs and eligible census-only HUs together geographically by their expected location on the ground (i.e., map spot numbers). Sometimes there is a large number of these census-only cases within a sample segment(s) in a BCU. If there is a large number (58 or more) of these census-only HUs in all the sample segments in the BCU, they are subsampled. Within sample segment(s) in a BCU, the census-only units are selected at a constant rate of 1-in-3.

If there are fewer than 58 eligible census-only HUs within all sample segments or there are fewer than 58 of these units within a BCU that was not subsampled during the P-sample selection, all of these eligible census-only HUs are included in the PES PI sample. We expect the number of census-only HUs in sample P-sample segments to be about 6 percent of the number of PI-sample HUs.

### 5.3 Identifying the E-sample HUs

The E sample is identified after the PES PI is conducted but before the Person Matching operations begin. However, the process will be discussed here to provide continuity in the sampling discussion.

The E sample contains HUs enumerated in the census in the same sample areas as the P sample. The source of the final set of census HUs is the Census Unedited File (CUF). [REDACTED]

One primary goal of E-sample identification is to geographically overlap the P and E samples to the extent possible. Having overlapping P and E samples reduces the E-sample PFU workload. Overlapping the P and E samples also improves the precision of the dual system estimates and the cost effectiveness of succeeding field operations.

The identification of E-sample HUs is done by mapping the final set of census HUs onto the previously-identified PES segments. The final census HUs are divided into two categories; those that are on the PI frame and those that are not. If a unit maps to a unit on the PI frame and is selected for PI, then it is in the E sample. Those units that did not map to the PI frame are Census Adds.

In addition, the census inventory of HUs may change between the IHU Matching and IHUFU operations and the identification of the E sample. For example, a unit may be added to the census address list through a census operation. The added unit could also have been on the census list at the time of IHU matching and followup, but declared erroneous for some reason such as

[REDACTED]

duplication. These erroneous units that are still in the census would also be treated as Census Adds for E-sample identification. These Census Adds are assigned to the already-formed segments geographically.<sup>17</sup> For the 2020 PES, Census Adds in sample segments will not be subsampled. The E sample is expected to contain about the same number of HUs as the P sample.

## **6.0 Person-based Operations**

Once the initial HU and sampling operations have been completed, a series of person-based operations is performed. These operations aim to create an independent roster of people who are living at each HU selected for the PES PI. The person-based operations include:

- PI.
- automated geocoding.
- automated Residence Status Coding.
- Clerical geocoding.
- Clerical Residence Status Coding.
- Person Computer Matching.
- Person BFU Clerical Matching.
- PFU.
- Person AFU Clerical Matching.

In the PES PI, a series of questions is asked to identify the Interview Day and Census Day residents of the sample address, and to determine where each of these people should have been counted for the census and all other places where the person could have been counted. During the Person Computer and BFU Clerical Matching operations, the information collected in the PI is compared to information collected in the census. Discrepancies and missing information are then followed up in PFU. The AFU Clerical Matching Operation uses the PFU results to resolve the cases.

### **6.1 Person Interview**

The PES PI will be conducted at the selected HUs to build an independent roster of people currently living at each HU. Emphasis will be placed on collecting information on the Census Day and Interview Day residence of people associated with the HUs (nonmovers and in-movers) based on residence rules. This method, by which all people who live at the sample address at the time of the PI are included in the P sample, is known as PES Procedure B. Information on the Census Day residence of these people is collected to facilitate matching. The 2020 PES will use Procedure B.

The 2010 CCM used Procedure B+. When the P sample also includes people who lived at the sample address on Census Day but who died after Census Day or whose Interview Day residence is out of scope for the coverage measurement universe (that is, has no probability of selection for the P sample), this is known as PES Procedure B+. An example of how the 2010 CCM employed

---

<sup>17</sup> Variables such as street name, house number, unit number, and rural route box are used.



the B+ methodology was collecting Census Day information on a person who was a Census Day resident of the sample address, but who moved into a college dorm between Census Day and Interview Day. The college dorm represents a residence that is out of scope for PES since it is a GQ.<sup>18</sup>

To maintain the independence needed to preserve the validity of the DSE methodology, the PI field operation should not overlap with the Nonresponse Followup (NRFU) operation within a BCU.

In the PI, the interviewers' primary focus is establishing the Census Day and Interview Day residence for people at the sample address at the time of interview, both nonmovers and in-movers since Census Day. Data will be collected for people (either individuals or entire households) about whom the respondent is knowledgeable and who have moved out since Census Day. Outmover data is collected to reduce the PFU workload and to obtain an interview closer to Census Day for these possible E-sample cases. This information also helps to identify people who may have been enumerated at more than one location in the census. Situations where all Census Day residents of an HU have moved out before Interview Day present a challenge for data collection and in resolving Census Day residence status.

For each household in the PI, the data collection process will include questions to establish where each person should have been counted in the census (their Census Day address) and information on where each person could have been counted, such as in-mover and alternate addresses. These in-mover and alternate addresses will be used to assist in determining the Census Day residence status of each person. They will also be used to identify possible erroneous enumerations (including possible duplicates).

The PI will be conducted using a computer-assisted person interview (CAPI) instrument on a laptop computer. Having an automated instrument allows the interview to be tailored to the specific living situation and demographic characteristics of all people, as determined by their responses. A series of roster probes will be asked to capture information about any other person who has stayed at the residence that the respondent did not previously mention. These questions ask about babies, foster children, relatives, friends, or someone else who may have stayed temporarily at the sample address.

After rostering, questions are asked to gather demographic information and additional addresses where a person could have been counted in the census. The questions are geared towards collecting addresses at any of the following places where a person might have resided (Shoemaker, 2008c):

- Census Day address other than the sample address.
- college address.

---

<sup>18</sup> In the 2010 CCM under Procedure B+, outmovers were included when they moved out from the sample address into a place that had no chance for selection in the P sample. These included people living in GQs or in places outside the U.S., people who experiencing homelessness, and people who died between Census Day and Interview Day.

- military service address.
- residence for a job.
- relative's address where the person may stay, including joint custody situations.
- seasonal home address.
- other GQ (nursing home, correctional facility, group home, etc.).
- any other place stayed often.

Questions are then asked as to how often a person “cycles” (goes back and forth) between the sample address and any alternate address. This method is similar to the one employed in the 2010 CCM PI and the 2000 A.C.E. person followup and evaluation followup interviews. Responses to these questions are needed so the clerical matching staff can apply the census residence criteria to determine where the person should be counted.

The PI quality assurance plan has the following components:

- All interviewers undergo online and classroom training.
- Interviewers who struggled in training are observed by field supervisors.
- Automated instrument edits to detect and correct common data-entry errors. Some edits prevent the interviewer from moving forward until a correction is made, while other edits warn the interviewer of a possible error but allow them to move on without a correction.
- A reinterview of selected cases, as described below.

The reinterview sampling is performed by the Sampling, Matching, Review and Coding System (SMaRCS). There are four ways a case can be selected for reinterview:

- Analytic, which selects cases that appear to be unusual. For example, a case that is completed very quickly is suspicious.
- Random, which selects cases using a systematic sampling approach.
- Supplemental, which is used when clerical resolution staff suspect an interviewer may not be following proper procedures and are looking for more information.
- Hard Fail, which is used when an interviewer is confirmed to have falsified work. In this situation, all of the interviewer's cases that have not already been selected for reinterview are added to the reinterview workload.

The reinterview is conducted by a separate staff of interviewers. As completed reinterview cases are transmitted from the field, the SMaRCS matches the PI and PI reinterview data to determine a match code. Reinterview cases that are not assigned a “Pass” or “reinterview noninterview” match code during computer matching go on to a clerical matching stage to determine a final match code. The reinterview program is designed to sample 15 percent of the PI cases.

## **6.2 Automated Geocoding**

The in-movers' Census Day addresses, and the alternate addresses provided by the respondents in PI, are geocoded. The geocodes are used during person matching to create search areas for person matching that encompass the BCU of the in-mover or alternate address and its surrounding BCUs. Since addresses are collected during both the PI and the PFU, geocoding will be required for addresses obtained by each operation at two different times. Respondent-provided in-mover and alternate addresses collected in the PI will be subjected to both automated and clerical



geocoding, while addresses collected in PFU will only be clerically geocoded (during the Person AFU Clerical Matching operation).

The results of automated geocoding for addresses collected in the PI will be input to the Clerical Geocoding operation. The Clerical Geocoding operation will attempt to geocode the addresses that could not be given a geocode during automated geocoding. The results from these two operations will be used to conduct Person Computer Matching and Person Clerical Matching, as well as Clerical Residence Status Coding. The inclusion of clerical geocoding results prior to computer matching is a change from 2010.

The Census Bureau's MAF/TIGER<sup>19</sup> database is the primary source of the geocodes. Geocoding can occur at various levels of geography. The most detailed geocoding occurs when the respondent-provided address matches to a specific record on the MAF. If the address matches to the MAF, we receive the matching MAF Identifier (MAFID) as well as the geocode. The least detailed geocode is assigning only a state. Several county codes can be assigned if only a city or town is provided and that city or town lies in more than one county. In general, geocoding is successful for matching purposes if the address can be assigned to a specific county and census BCU or a small group of census BCUs.

If the automated geocoding is not successful in determining the MAFID or the BCU of the respondent-provided address, then the address is provided to Clerical Geocoding for the clerical matchers to attempt to geocode using additional information collected during PI.

### **6.3 Automated Residence Status Coding**

Each person rostered in a PI HU that is coded as a complete or partial interview is assigned a residence status code. This includes people in census-only HUs selected to receive the PI. These codes ultimately are used to indicate whether a person should be included in the P sample or not. A person rostered in a census-only HU is assigned a residence status code but is not eligible for the P-sample. The residence status codes are assigned by computer following the PI and can be reviewed and reassigned during Clerical Residence Status Coding or any part of person clerical matching based on information provided in PI or PFU.

The residence status codes for P-sample people are as follows:

- nonmover (N)
- inmover (I)
- unclassified (U)

The PI is designed to roster all people at the sampled HU on Interview Day. If the respondent can provide information on people who lived at the sample HU on Census Day, but no longer do, then this information is collected. As such, some of the people rostered may not turn out to be

---

<sup>19</sup> Master Address File/Topologically Integrated Geographic Encoding and Referencing System. The MAF/TIGER System serves as the national repository for all of the spatial data and residential address data needed for the Census Bureau's data collection, data tabulation, data dissemination, geocoding services and map production.

residents of the sample address on Interview Day and would therefore not be included in the P sample. The residence status codes for these non-P-sample people are as follows:

- outmover (O).
- out of scope (X).

If the computer determines that a clerical review is needed to determine the residence status, the person is given a residence status code of “R” which will flag the record for clerical review.

#### **6.4 Person Matching**

The person matching design for the 2020 PES is similar to the 2010 person matching design.

The Person Matching operations include two parts: Person Computer Matching and Person Clerical Matching. The Person Computer Matching operation establishes matches when straightforward relationships between records exist. The Person Before Followup (BFU) Clerical Matching operation allows for trained clerical matchers to review the computer matches and to determine match status for cases that were possibly matched or not matched by the computer. Clerical matchers will search for matches and duplicates in appropriate search areas, using various tools and methods made available to them in the clerical matching software.

The Person Clerical Matching operation will be used to search for matches between PI person records and census enumerations within each PES sample BCU and surrounding BCUs (this is called the “sample BCU search area”). It will also be used to search for census enumerations of in-movers in the BCU containing their Census Day address and in the surrounding BCUs (this is called the “in-mover address search area”). The in-mover’s Census Day address will typically be outside the BCU containing the sample address. A clerical search for census person enumerations where alternate address information was provided in PI or PFU will also be conducted in the BCU containing that address and in the surrounding BCUs (called the “alternate address search area”). For possible matches and possible duplicates found in the nationwide computer search, the clerical search will be conducted only at the linked unit located outside of the sample BCU search area (that is, the search area is limited to that specific unit).

Person matching is designed to:

- determine which PI person records in P-sample HUs belong or might belong in the P sample.
- determine whether the PI person records and E-sample person records have sufficient information for followup.
- determine PI person records that are duplicated within the sample BCU search area.
- determine E-sample person records that are a duplicate of a census enumeration within the sample BCU search area or throughout the U.S.
- determine PI in-mover records that link to a census enumeration within the in-mover address search area.
- determine PI person records that link to census enumerations within the alternate address search area.



- determine PI person records that link to census enumerations at the addresses where computer matching identified matches or possible matches during the nationwide computer search.
- identify cases that require more information to get the match status, enumeration status, and residence status that will be sent to PFU.
- use the results of the PI and PFU interviews to determine match status, residence status (PI people), and enumeration status (E-sample people).

The Person Matching operations compare people in the PI- or E-sample person enumerations to all census person enumerations and identify records referring to the same person. Correct enumerations and duplicates are determined through matching PI enumerations to the list of all census enumerations. If an E-sample person matches to the same person as another census enumeration, it means the person was duplicated by the census.

After the Person Computer Matching and the Person BFU Clerical Matching operations are completed, the PFU operation provides more information for cases with unclear or missing match status, residence status, or enumeration status. Cases with missing demographic information are not sent to PFU if the relevant statuses have been determined. After PFU collects more information, a Person After Followup (AFU) Clerical Matching operation takes place to use the supplemental information obtained via the PFU interviews.

Sections 7.5 through 7.8 and Section 7.10 provide more details on the Person Matching operations.

## **6.5 Person Computer Matching**

The Person Computer Matching involves three different searches:

- PI to census - PI people are matched to the entire census to find matches both within the sample BCU search area and throughout the census.
- E sample to census - E-sample people are matched to the entire census to find duplicates both within the sample BCU search area and throughout the census.
- PI to PI within-BCU - PI people within the BCU are matched to find within-BCU PI duplicates only.

For each search outlined above, Person Computer Matching involves several steps:

1. Preparing the Data - Census and PI data fields used in the computer matching algorithm need to be consistent. They include data items such as gender and computed age. In addition, the phone number, a key component in the matching system, is edited. Finally, names are prepared by removing punctuation, titles (Mr., Mrs., Dr., etc.), and special characters (including Spanish characters). A name variant list is used to standardize first names (for instance "Cathy" and "Catherine" are considered the same).
2. Computer Matching - Person records are compared across all living quarters outside of their own (i.e., a person cannot match within their own HU or GQ). A person can link to

multiple people but to only one person per living quarters. Matching is accomplished in two steps. The first step matches all people across the universe. The second step builds on the links established in the first step but allows additional flexibility for finding matches within other households.

3. Geographic Assignment - Each linked pair is assigned a geographic code indicating the relationship of the link. The codes are:
  - within the sample BCU.
  - in a surrounding BCU.
  - in county but outside the sample BCU and surrounding BCUs.
  - in a different county in the same state.
  - in a different state.
4. Computer Modeling - Modeling is used to establish which links should be considered matches, possible matches, or not a match. Rules are applied based on geographic proximity (see Geographic Assignment above), household patterns, phone numbers, person characteristics, and type of living quarters (HU or GQ). There are two types of modeling. The first is within-household modeling, which takes the household structure into consideration. The second type is residual modeling, which matches individuals without the benefit of the household structure.
5. Creating the Output - The output from computer matching contains identifying information for each person record in a linked pair and the match code for that pair (either match or possible match).

Each of the three searches uses the same software and the same steps in matching. For the PI-to-census search, the data are matched only once but there are three separate modeling processes. Links in the sample address search area or in the nationwide search area are modeled with respect to the sample address. Links in the in-mover address search area are modeled with respect to the in-mover address. Links in the alternate address search areas are modeled with respect to the alternate address. For each of the modeling steps, the geographic distance category is reassigned based on the distance of the address being modeled. This adjustment can make it more likely to call a link a match when modeling with respect to a different address. In general, links with closer geographic proximity are more likely to be called matches.

## **6.6 Person Before Followup Clerical Operations**

There are three clerical operations associated with person matching prior to PFU:

- Clerical Geocoding.
- Clerical Residence Status Coding.
- Person BFU Clerical Matching.

Clerical Geocoding occurs after the Automated Geocoding and before the Person Computer Matching. The Clerical Residence Status Coding occurs after the Clerical Geocoding and Automated Residence Status Coding. The Clerical Residence Status Coding may occur parallel



to Person Computer Matching. The Person BFU Clerical matching occurs after the Clerical Residence Status Coding and after the Person Computer Matching.

To maintain and check the quality of the Clerical Geocoding and Clerical Residence Status Coding stages, the analysts will review a sample of work from every technician. They will also review specific records referred to them by the technicians. We estimate that 25 percent of addresses will be subject to verification with a target AOQL of 6.13 percent for each stage.

After training, technicians will start production work in 100% QC status. If a technician is performing poorly, the technician may be required to undergo retraining. Technicians that perform poorly even after retraining will be removed from the operation.

To help ensure the quality of the matching and ensure the newer analysts understand procedures, the experienced analysts will informally review the work of the newer analysts.

### **6.6.1 Clerical Geocoding**

After automated geocoding, all cases with ungeocoded inmover or ungeocoded alternate addresses will be reviewed during the Clerical Geocoding Technician stage.

Any inmover or alternate address that is not geocoded to a BCU by automated geocoding will be reviewed by the technicians. They will have access to census maps for the entire country, the MAF browser, the online ZIP+4 directory, and the Internet to attempt to geocode the address. They will also have access to other information collected in the PI such as landmarks, cross streets, names of other household members, and names of nearby neighbors.

Clerical matchers also review the PI data and geocode(s) for each inmover addresses to determine how confident they are that the census person would be found in the inmover address search area, given that the person was correctly enumerated in the census. Clerical matchers assign an inmover address code to each inmover address, based on this review.

### **6.6.2 Clerical Residence Status Coding**

After automated residence status coding, PI cases with a “review” residence status code from automated residence status coding will be reviewed during the Clerical Residence Status Technician stage. In some cases, the computer cannot assign a residence status code to a person rostered in the PI, but notes collected during the PI may help resolve the status. In these cases, a technician will review the data collected in the PI and will assign a residence status code. The Clerical Residence Status Technician stage within a BCU occurs after the BCU completes Clerical Geocoding, since some cases may require geocodes to determine the correct residence status code.

### 6.6.3 Person Before Followup Clerical Matching

During the BFU Technician stage, technicians will review the results of Person Computer Matching. In particular, technicians attempt to match PI records to the census, find duplicates in the PI, and find E-sample duplicates.

The technicians will review each E-sample case with insufficient information for followup to attempt to update the record with information from other census data (including electronic images of census forms). An example would be adding the last name of a parent listed on a census form to children on the form whose last names are missing. PI cases with incomplete names or characteristics will be updated using other PI data.

As part of the Person BFA stage, analysts independently review every technicians' work and resolve any cases that the technicians could not resolve. The analysts will perform a full review of BCUs selected for the person matching QC sample. They will also review specific person records sent by the technicians in Person BFT stage. It is expected that about 25 percent of all BCUs requiring person matching will be subject to this QC operation, with an expected AOQL of 3.5 percent.

After training, technicians will start production work in 100% QC status. If a technician is performing poorly, the technician may be required to undergo retraining. Technicians that perform poorly even after retraining will be removed from the operation.

To help ensure the quality of the matching and ensure the newer analysts understand procedures, the experienced analysts will informally review the work of the newer analysts.

After this initial matching operation, all E-sample nonmatches, selected PI nonmatches, and all possible matches and duplicates are followed up in the field via the PFU operation.

## 6.7 Person Followup

Person Followup cases are identified after the Person BFU Clerical Matching operation is completed. The PFU interview attempts to collect additional information needed to establish Census Day residence for P-sample cases. It also attempts to resolve enumeration status for nonmatched E-sample cases. People coded as having insufficient information for followup are not sent to PFU. The types of cases sent to PFU are:

- Possible matches (P- or E- sample), not including possible matches identified during nationwide computer matching.
- P-sample partial-household nonmatches.
- P-sample whole-household nonmatches with a proxy respondent in the PI and either no people in the matching census unit or no matching census unit.
- P-sample whole-household nonmatches in BCUs with a high rate of P-sample nonmatches (more than 45 percent).
- P-sample whole-household nonmatches where the PI interviewer changed the sample address.
- Inmovers with ungeocoded inmover addresses (regardless of match status).



- Conflicting households which occur when addresses that matched during IHU operations have completely different rosters in the census and PI (nonmovers).
- Unclassified P- or E-sample people with sufficient information for followup, who are not duplicates, regardless of match status.
- P-sample people who are identified as match or possible match in the nationwide computer matching (followup at sample address).
- People identified in the nationwide search who match or possibly match to P-sample people (followup at the nationwide address) by personal visits.
- E-sample nonmatches.<sup>20</sup>
- E-sample whole-household nonmatches whose addresses were not included in Initial HU Matching and flagged for a geocoding check.
- E-sample people who are duplicates or possible duplicates to a nationwide person (followup at the sample address).
- People identified in the nationwide search who are duplicates or possible duplicates to other E-sample people (followup at the nationwide address by personal visit).
- Any case sent to PFU by an analyst.
- P-sample possible duplicates within the sample BCU.
- E-sample possible duplicates within the sample BCU or the surrounding BCUs (at least one of the pair is an E-sample person).
- PI Inmovers who are not found at their respondent-provided Census Day address (i.e., not matched at their in-mover address), where the PI data was collected from a proxy respondent.
- Nonmovers at addresses where the IHU operations determined that no HU existed at the time of followup, but the potential existed for an HU at the time of PI.
- E-sample matches to outmovers at addresses where the IHU operations determined that no HU existed at the time of followup.

Note that a case can be sent to PFU for multiple reasons. The PFU data collection methods will include establishing where the person should have been counted in the census (their Census Day residence) and collecting information on alternate locations and addresses where the person could have been counted on Census Day.

The PFU interview will be conducted using a docuprinted paper questionnaire. This method allows the PFU form to be tailored to collect only the information needed for each person. The paper questionnaire will be used for final coding of the Person Matching results.

The PFU questions are geared towards collecting addresses at any of the following places where a person might have resided:

- Census Day address other than the sample address.
- college address.
- military service address.
- residence for a job.

---

<sup>20</sup> People enumerated by the census in Transitory Locations are not eligible for PFU.

- relative's address where the person may stay, including joint custody.
- seasonal home address.
- any other place stayed often.
- other GQs (nursing home, correctional facility, group home, etc.).

To assign a residence status code, the PFU interview then asks for dates spent at each residence and includes some semi-scripted probes to help matching staff determine at which address the person should be counted in the census according to census residence criteria.

In addition, a followup on matches, duplicates, possible matches, and possible duplicates identified in the nationwide computer match will be performed. The part of the link that is located outside PES sample and surrounding BCUs (referred to as the "nationwide" address) will be followed up by personal visits as part of the nationwide PFU operation. The part of the link located inside the search area will be followed up in the field as part of the regular PFU operation.

The PFU quality assurance plan has the following components:

- All interviewers undergo online and classroom training.
- Interviewers who struggled in training are observed by field supervisors.
- Field supervisors review and edit completed PFU forms for accuracy and completeness.
- Each PFU questionnaire undergoes an office edit.
- Reinterview, described below.
- Following reinterview, each case is edited by a field supervisor and then office staff.

The purpose of the reinterview is to ensure that the original interviewer contacted the PFU case. Reinterviews are chosen by random or supplemental sampling. Once a case is selected for reinterview, a reinterviewer will visit the followup address or make contact by telephone and conduct the reinterview with the same person who responded to the PFU interview to confirm that the PFU interview respondent was contacted by the interviewer. About 15 percent of each interviewer's completed eligible cases are randomly selected for reinterview.

## **6.8 Person After Followup Clerical Matching**

Following the PFU interview, Person AFU Clerical Matching operations are performed. The Person AFU Clerical Matching consists of four stages:

- AFU Technician.
- AFU Analyst.
- AFU BCU Review.
- AFU Outlier Review.

As PFU forms are returned from interviewing, they are assembled into batches to be reviewed by the clerical matching staff. First, these batches are reviewed by technicians in the Person AFT stage. Then, during Person AFU Analyst, analysts perform a review of the technicians' work selected for QC purposes. The analysts also resolve any cases that the technicians could not resolve. This process is similar to the process for all other PES clerical matching operations.



About 20 percent of all PFU forms will be reviewed for QC purposes. The target AOQL is 6.13 percent.

After training, technicians will start production work in 100% QC status. If a technician is performing poorly, the technician may be required to undergo retraining. Technicians that perform poorly even after retraining will be removed from the operation.

To help ensure the quality of the matching and ensure the newer analysts understand procedures, the experienced analysts will informally review the work of the newer analysts.

During the AFU Technician and AFU Analyst stages, the PFU interview results are clerically reviewed and the people are coded according to a pre-determined set of matching codes. Technicians and analysts are allowed to make corrections to previous coding assignments if errors are discovered from the Person BFU Clerical Matching operations, even if the people were not included in PFU. These match codes provide the basic information needed for measuring both net coverage error and the components of census coverage. This clerical review also includes assignment of residence status codes and the clerical geocoding of any additional respondent-provided addresses collected in the PFU interview.

All PES sample BCUs are eligible for AFU BCU Review and AFU Outlier Review stages. BCUs are flagged for AFU BCU Review by technicians or analysts in the AFU Technician or AFU Analyst stage when certain conditions make it necessary to review the BCU as a whole instead of the batched forms. AFU Outlier Review stage is an analyst-only stage that targets the BCUs that may need further review. Analysts will review the BCU and write journals to explain any problems with the BCU. Analysts will also recode any cases that may be incorrectly coded. The journals are useful during analysis when particular BCUs stand out. BCUs are flagged for outlier review by Headquarters staff or by having a high outlier priority score. The outlier priority score is computed as the weighted sum of cases that do not match in the net error context in the P sample and cases in the E sample that do not link back to the P sample or are considered erroneous enumerations for net error.

Person AFU Clerical Matching is the final PES person operation. Output files with the results from the PI and person matching operations will be made available for the PES Estimation activities (see Section 9). People that remain unresolved will be handled through statistical techniques for missing data.

## **7.0 FHU Operations**

The primary objective of the FHU operations is to process changes to the HU inventory that occurred between the preliminary census universe (the Decennial Master Address File) used in the IHU matching and the final census HU universe (the Census Unedited File). Importantly, the FHU processes the census adds and deletes, allowing proper evaluation of the final census results.

The PES FHU operations are the

- FHU Computer Matching and Processing.
- FHU BFU Clerical Matching.
- FHU Followup.
- FHU AFU Clerical Matching operations.

As a whole, these four operations work together to

- identify geocoding errors and any remaining HU duplicates.
- update pre-existing files.
- resolve match, enumeration status, and HU status for all P- and E-sample HUs not resolved in prior PES operations.

## 7.1 FHU Computer Matching and Processing

In the FHU Computer Matching and Processing, information from all previous stages of computer and clerical matching from both HU and person procedures, coupled with final census data, is used to update HU records. Using the person matching results, the Computer Processing assigns geocoding error codes to P- and E-sample units where the people are coded as such, and assigns Census Day HU status to unresolved HUs based upon specific PI results. These PI results include vacant HUs and records that were not HUs on Census Day.

During FHU Computer Matching, some addresses listed in the PES sample BCUs during IL are computer matched to the CUF addresses (of HUs and GQs) within the sample BCUs and one ring of surrounding BCUs. This matching operation may identify:

- matches and possible matches between the IL addresses and census addresses.
- potential PES IL duplicates within the sample BCU.
- potential census duplicates.

In addition, FHU Computer Processing identifies units that will go to FHU BFU Clerical Matching. Computer Processing will assign a work flag to the following types of units:

- P-sample links to census deletes.<sup>21</sup>
- census adds<sup>22</sup> that are in the sample BCU (including subsampled out).
- census adds that are in the surrounding BCUs to the PES sample BCU.
- E-sample units that were duplicates where the primary<sup>23</sup> was deleted from the census address canvassing results file and is not on the CUF.
- E-sample units that linked to a unit that is no longer in the P sample.

<sup>21</sup> A census delete is a HU on the preliminary census address list used for Initial HU Matching, but NOT on the CUF.

<sup>22</sup> A census add is a HU NOT on the preliminary census address list, but is on the CUF.

<sup>23</sup> If two or more records within the same system (PES or census) represent the same person or living quarters, one record is designated as the primary and given an appropriate non-duplicate match code; the other record(s) is(are) designated as the duplicate(s) and given an appropriate duplicate match code.



- any P- or E-sample unit (not covered above) for which HU Computer or Clerical Matching has not yet been conducted.<sup>24</sup>

Lastly, Computer Processing identifies PES and census units, including census GQs, with links to census units not on the CUF, and then unlinks those units. Once the data files have been updated and the units flagged, the next process can begin--the FHU BFU Clerical Matching.

## 7.2 FHU Before Followup Clerical Matching

In the FHU BFU Clerical Matching operation, files containing the entire BCUs including the flagged HUs from Computer Processing and Matching are made available for matchers to view. Matching staff will attempt to match by address, HU description, or map spot location, all P-sample units and census units that were flagged earlier. Clerical matchers will have access to data from all previous PES operations, including maps. Clerical matches will also have access to map-spotted maps from the census corresponding to the CUF data. Those units that remain unresolved will be flagged for FHUFU, unless previous attempts were made to followup the units during IHUFU. The FHU BFU Clerical Matching operation will have multiple sources of information including:

- The data prepared for FHU BFU Clerical Matching in FHU Computer Processing and Matching.
- completed IHUFU forms.
- completed PFU forms.
- PES Person AFU Clerical Matching results.
- Census maps with map spots for HUs and GQs of the same vintage as the CUF.
- PES maps with map spots and updates from IL and IHU Matching and Followup.

In addition to identifying matches between PES and census units, matching staff will attempt to identify duplicate units:

- between PES P-sample units within the PES sample BCUs.
- between E-sample units and other census units in the PES sample BCU and in its surrounding BCUs.

Clerical matchers will search all the census addresses added in the sample BCU and surrounding BCUs to identify duplicates to the E-sample addresses.

Also, analysts are allowed to change match codes previously assigned if errors are uncovered, and may attempt to resolve the case.

---

<sup>24</sup> All P-sample HUs should have been processed through Initial HU Computer or Clerical Matching by this point. They are included here to handle unexpected problems that might prevent a BCU from going through the Initial HU Matching system.

The FHU BFU operation will include BFU Technician, BFU Analyst, and quality control components, similar to IHU BFU.

In the Final Housing Unit Before Followup Analyst (FHU BFA) stage, analysts perform the QC of the technicians' work and review cases referred by the technician. The QC process for clerical matching consists of analysts performing a full review of block clusters selected for the matching QC sample. For the FHU BFU Clerical Matching operation, approximately 33 percent of all block clusters going through the operation will be checked. The target AOQL is 3.5 percent.

After training, technicians will start production work in 100% QC status. If a technician is performing poorly, the technician may be required to undergo retraining. Technicians that perform poorly even after retraining will be removed from the operation.

To help ensure the quality of the matching and ensure the newer analysts understand procedures, the experienced analysts will informally review the work of the newer analysts.

### **7.3 FHU Followup**

The FHUFU is a field operation performed by PES field representatives to resolve questions, discrepancies, and missing data from the FHU BFU operation. In making assignments, care is taken to observe the independence rules followed for all PES operations. Field representatives will use a preprinted paper form along with PES maps updated with the results from the PES IL and IHU operations with map spots and the census maps to locate flagged HUs.

The questions on the FHUFU forms will be tailored for each specific case given its situation. The following information will be collected:

- BCU number (for a unit that could be a duplicate in a surrounding BCU).
- Evidence to determine if there was a HU at the address on the date of the followup visit. If the case in question is not a HU, a statement should be provided as to the reason for this determination.
- Whether or not two or more units identified as possible matches are the same unit.
- Whether or not two or more units identified as possible duplicates are the same unit.
- HU status updates for certain situations.
- Response to case-specific questions provided by the clerical matchers.

The questions will be the same as those on the PES IHUFU form for similar types of followup situations, except for changes needed due to the different reference dates for the initial and final followup interviews.<sup>25</sup>

Interviewers will use the reference list<sup>26</sup> provided during the FHUFU to determine if any units sent to followup match other units on the reference list. All of the information collected during followup will be used to assist in determining the status of followed up HUs as of Census Day.

---

<sup>25</sup> If any serious problems are encountered with the IHUFU form, the FHUFU form may be changed.



The FHUFU quality assurance plan has the following components:

- All listers undergo online and classroom training.
- Listers who struggled in training are observed by their field supervisor.
- Field supervisors review and edit completed forms for accuracy and completeness.
- A separate staff of listers, called QC listers, perform a check of data accuracy for randomly selected FHUFU forms in each BCU. Approximately 15 percent of all cases sent to FHUFU will be checked in the field, with a target AOQL of 4.0 percent.
- A field supervisor edits the QC forms for accuracy and completeness.
- Each BCU undergoes an office edit.

Note, the FHUFU quality assurance plan is identical to the IHUFU quality assurance plan.

#### **7.4 FHU After Followup Clerical Matching**

Like the IHU AFU Clerical Matching, the FHU AFU Clerical Matching contains FHU AFU Technician, FHU AFU Analyst, and FHU AFU Outlier Review components.

The FHU AFU Clerical Matching operation will have multiple sources of information to help resolve HU status, match status, or enumeration status such as:

- data used for PES FHU BFU Clerical Matching.
- completed FHUFU and IHUFU forms.
- completed PFU forms.
- PES FHU BFU Clerical Matching results.
- PES Person AFU Clerical Matching results.
- PES maps with any updates from FHUFU.

Using this wealth of data, matching staff will attempt to determine unit status, match status, or enumeration status and assign appropriate codes to each case sent to FHUFU. An output file with the results from the PES HU operations will be made available for the PES Estimation activities. HUs that remain unresolved will be handled using statistical techniques for missing data.

In the FHU AFU Analyst stage, analysts perform both production and QC; unlike the other clerical operations, there are no technicians. The QC process for clerical matching consists of analysts performing a full review of block clusters selected for the matching QC sample. For the FHU BFU Clerical Matching operation, approximately 33 percent of all block clusters going through the operation will be checked. The target AOQL is 3.5 percent.

To help ensure the quality of the matching and ensure the newer analysts understand procedures, the experienced analysts will informally review the work of the newer analysts.

---

<sup>26</sup> The FHU reference list is intended to be a list of all PES and census HUs in the PES sample BCU and its surrounding BCUs.

## 8.0 Estimation

All of the details included in this section are accurate as of the date of this memo. However, various aspects of the estimation process are subject to change.

The PES Estimation process contains several operations, which will ultimately lead to the production of estimates of coverage for both HUs and people in HUs (hereafter referred to as “people”). This includes estimates of net coverage as well as estimates of the components of census coverage.

Like the 1990 Post-Enumeration Survey, the 2000 Accuracy and Coverage Evaluation (A.C.E.), and the 2010 Census Coverage Measurement (CCM) Survey, the 2020 PES will estimate net coverage by using dual system estimation (DSE) to generate the population estimates of HUs and people. Similar to the 2010 CCM, we will use logistic regression modeling instead of the traditional post-stratification to estimate the parameters in the DSE formula. Logistic regression modeling will also be used in some of the missing data operations; they will be described in more detail later.

For estimates of correct and erroneous enumerations, we will use a more straightforward estimator – a basic summation of ratio-adjusted survey weights. The adjustment ensures that the estimates of correct and erroneous enumerations add to the appropriate number of data-defined census enumerations for a limited set of groups.<sup>27</sup> Also, the ratio adjustment will typically reduce the variance of the estimates.

### 8.1 Missing Data for P-Sample People

There are four adjustments designed to account for missing data for people in the P sample:

- imputing missing demographic characteristics.
- applying a household-level noninterview adjustment for the PES Person Interview.
- imputing the probability that a person should be included in the P sample.
- imputing a match status.

---

<sup>27</sup> To estimate the components of census coverage, design-based estimates will be used for correct and erroneous enumerations. Estimates of omissions will be a combination of design-based and model-based.



### **8.1.1 Imputing Missing Demographic Characteristics for People**

Production of PES coverage estimates requires data on:

- Age.
- Sex.
- Tenure (owner versus renter).
- Relationship to the householder.
- Race.
- Hispanic origin,

to classify respondents by these important demographic characteristics, so they will be imputed whenever the data was not collected. We will impute these characteristics using the Census Edit and Imputation programs.

### **8.1.2 Person-Level Noninterview Adjustment for PES Household Noninterviews**

The PES accounts for a potential bias introduced because of households that were not interviewed in the PI. The units with “missing” interviews have sample weights associated with them. The noninterview adjustment procedure accounts for these cases by assigning a weight of zero to noninterviews and distributing their weight uniformly among interviewed cases with similar characteristics and response propensities. Weights from noninterviewed American Indian Country (AIC) HUs will be distributed only to other AIC HUs, and weights from noninterviewed non-AIC HUs will be distributed only to other non-AIC HUs. The noninterview adjustment is applied to people in interviewed HUs, for the purposes of dual-system estimation for people. Weights used in HU estimation are unaffected because all household responses are obtained by observation, if a respondent does not provide them.

A weight trimming procedure is applied after the missing data operations have been performed, but before the final coverage estimates are created. Trimming the weights reduces the likelihood that a small number of BCUs will have a large influence on the coverage estimates and their standard errors.

### **8.1.3 Imputing the Probability of P-Sample Inclusion for People**

The PES interview needs to determine if a person should be included in the P sample. The P sample of people includes residents of the P-sample HU on Interview Day.

We will assign a probability of P-sample inclusion status for all sample cases. Resolved cases that should be included in the P sample will receive a probability of 1. Resolved cases that should not be included will receive a probability of 0. For the unresolved cases, the PES uses an imputation procedure with logistic regression. Since the overstatement of the P-sample total was an error for the A.C.E., the PES Estimation Team will closely examine and review how the imputation handles these cases.

#### 8.1.4 Imputing the Probability of Match Status for People

The PES needs to determine if the P-sample person matched to a census enumeration in the net error search area.<sup>28</sup> This involves accounting for P-sample cases that may have an unresolved match status or an unresolved mover status. A record may have either or both of these statuses unresolved. This section lays out the methodology to account for both statuses and how the final match probability for net error estimation is assigned.

We will assign a match probability for all sample cases. Resolved matches will receive a probability of 1. Resolved nonmatches will receive a probability of 0. Cases with unresolved status need to have a match probability assigned to them.

Since both the match status and mover status can be unresolved, we will assign an overall match probability for unresolved cases based on the following conditional probability of mover status.

We first need to account for unresolved inmover status. The cases used for this imputation will be the resolved nonmovers and resolved inmovers. Using the resolved cases, we will determine the weighted proportion of cases that were nonmovers or inmovers. This will be done by forming cells using some of the covariates in the logistic regression models, such as Before Followup (BFU) groupings. This will allow the two mover probabilities – of being a nonmover and of being an inmover – to be assigned to each case within each grouping. (The probability of being a nonmover and the probability of being an inmover sum to one.) The weighted proportions will be assigned as these probabilities for the unresolved mover status cases.

We will then assign the probabilities of a case being a match given that it is either a nonmover or an inmover. When these conditional probabilities are unresolved, we will use logistic regression methods to impute them. These models and the general prediction equation will be used to determine the conditional probability of cases matching.

The conditionally predicted match rate for each unresolved match status case is then obtained. If the case is a nonmover, then the correct search area is the BCU and ring of surrounding BCUs. We first use the clerical matching information to see if matching staff were able to match the case in the BCU search area. If so, we will assign this conditional probability of a match for this case equal to 1. The clerical matching information may indicate that the case was a nonmatch in the BCU search area. If so, we will assign this conditional probability of a match equal to 0. If the matching results are unresolved, then we will assign the conditional probability of a match by using logistic regression.

The predicted match rate for each unresolved P-sample case given that it is an inmover is obtained next. One difference from the conditional match probability given that the case is a nonmover is that the correct search area for inmovers needs to be identified based on the PES interview information. The clerical results may indicate that a case links to a census enumeration outside the BCU search area but we cannot assume that it is the correct search area. For these cases the conditional match prediction will be assigned using logistic regression. Recall that

---

<sup>28</sup> The net error search area is typically the PES sample BCU and one ring of surrounding BCUs.



when assigning the predicted match rate for each unresolved P-sample case given that it is a nonmover, we treated cases with an unresolved mover status as nonmovers.

After a) the probabilities of the case being either a nonmover or inmover, and b) the conditional match probabilities of nonmovers and inmovers are estimated, we can estimate the probability of the unresolved case being a match.

## **8.2 Missing Data for E-Sample People**

There are two missing data procedures needed for people in the E sample: a procedure for imputing missing demographic characteristics and one for imputing enumeration status (correct or erroneous). However, imputation of enumeration status needs to be considered separately depending on whether we are estimating net error or components of census coverage. Each of these processes will be discussed in more detail in the following sections.

### **8.2.1 Missing Demographic Characteristics for E-Sample People**

For people in the E sample with missing demographic characteristics, we will not develop any special imputation procedures. We will match those records to the Census Edited File and obtain the imputed characteristics from that file based on the Census Edit and Imputation methodology.

### **8.2.2 Imputing Missing Enumeration Status for People in the E Sample (Net Error)**

The PES needs to determine if the census enumeration in the E sample was correctly enumerated in the net error search area. The probability of correct enumeration will be assigned for all E-sample cases. The cases with unresolved status will have a probability imputed for them.

As was done in the 2010 CCM, a logistic regression model will be used to generate the predicted probability of unresolved cases being correct enumerations. The independent variables in the logistic regression models used to predict correct enumeration probability will include matching and PES interview information that has been identified as good discriminators of enumeration status that are only available for cases in the E sample and not available for the entire census. The models will include, at a minimum, the main effects of the independent variables used in the regression models to determine the predicted probabilities of being data-defined, a correct enumeration, or a match in the overall dual system estimation formula, as noted in Section 8.3.

### 8.2.3 Imputing Missing Enumeration Status for People in the E Sample (Components of Coverage)

The definition of correct enumeration used for net coverage estimation attempts to minimize matching error. This definition is not appropriate for estimating coverage components, since it results in inflated estimates of erroneous enumerations. Therefore, some cases treated as erroneous enumerations for net error estimation are considered to be correct enumerations for component estimation. One example is that an enumeration may be treated as erroneous for net error because it was enumerated in the wrong location but, if the person was enumerated once and only once, then that enumeration would be correct for national estimates of components.

Another example is the treatment of cases with insufficient information. If they are data-defined, but do not have adequate information for PES followup, they are treated as resolved erroneous enumerations for net coverage. However, for component estimation, the PES will expand its matching operations to try to determine the enumeration status of these cases. If this matching effort is successful, some of the cases treated as erroneous enumerations for net error will be correct enumerations for estimating components. The ones that cannot be matched are unresolved for estimating components of coverage and require imputation.

As was done in the 2010 CCM, we will use a cell mean imputation methodology to assign the enumeration status for the unresolved cases. The A.C.E. used this methodology for the missing data adjustments of unresolved enumeration status for net error estimation in 2000. The enumeration statuses are in Table 5. The weighted averages of each type of enumeration listed in Table 5 will be estimated using the resolved cases and those weighted averages will be assigned to the unresolved E-sample cases.

**Table 5. Probabilities for PES Component Estimation**

National Definition	Subnational Information for Correct
Correct	In the BCU search area: p(1)
	Same County, Same Place: p(2)
	Same County, Different Place: p(3)
	Different County, Same State; Same Place: p(4)
	Different County, Same State; Different Place: p(5)
	Different State: p(6)
Erroneous	Duplicate: p(7)
	Other Reasons: p(8)

Note: Erroneous cases are erroneous for all geographic entities.

### 8.3 Person Net Coverage Error

There is a long history of using DSE in measuring census coverage. The DSE for people is calculated using the following equation:

$$DSE = \sum_{j \in C} \pi_{dd(j)} \times \frac{\pi_{ce(j)}}{\pi_{m(j)}} \times CB_j$$

where  $j$  indexes person records in the census. The predicted data-defined, correct enumeration, and match probabilities ( $\pi_{dd}$ ,  $\pi_{ce}$ ,  $\pi_m$ , respectively) will be obtained through logistic regression



modeling. In 2010, the correlation bias adjustment factor (CB<sub>j</sub>) was computed using sex ratios by race from the Census Bureau's Demographic Analysis program. This adjustment accounted for differences between the PES target population and the census universe, namely removing people in group quarters (GQs) and in remote Alaska. Correlation bias adjustment factors are currently being researched for young children and Hispanic Origin.

We will then estimate net coverage by comparing the estimate of the true population (from the DSE) to the census count.

## **8.4 Person Components of Census Coverage**

For people in HUs, estimates of correct and erroneous enumerations will be formed at the national level and for various subpopulations (or estimation domains) using data from the E sample. The requirements for an enumeration to be considered correct for component estimation are less stringent than those treated as correct for net coverage. See Table 5 for the components of coverage that are correct for component estimation but not DSE. We will tally and report the total number of census whole-person imputations and omissions overall and for various sub-categories. Census whole-person imputations are tallied from census files, and omissions are calculated by deduction.

### **8.4.1 Calibration Adjustment for E-Sample People**

Estimates of correct and erroneous enumerations will be formed by inflating ratio-adjusted survey weights attached to person records from the PES E sample in the particular estimation domain.

A set of pre-defined cells is formed, and the sampling weights for the E-sample cases in each cell are ratio adjusted so that the sum of the weights equals the census totals for those cells. These cells may be based on race/ethnicity, tenure, age/gender groupings or other demographic variables. Geographic variables, such as region and state, can also be used, and the census totals are the number of data-defined census person records in each cell. Depending on sample sizes and the size of the adjustment factors, some collapsing of the original cells may be necessary.

### **8.4.2 Erroneous and Correct Enumerations for People**

For component estimation, the PES will estimate the total number of erroneous enumerations. The PES will also estimate erroneous enumerations by type:

- Duplicate person enumerations.
- People who should not have been enumerated at all (erroneous for all other reasons).

The PES will use the following definition of being a correct enumeration at the national level when evaluating the components of census coverage of people in the 2020 Census:

The person enumeration is considered to be correct if the record corresponds to a person that should have been included anywhere in the coverage universe<sup>29</sup>. If such a person was included multiple times, one of the enumerations will be considered correct and the other enumeration(s) will be erroneous.

As stated previously, there are stricter requirements to be a correct enumeration for net coverage than for the estimation of the components of census coverage.

For the estimation of correct enumerations as a component of census coverage at the state level, an enumeration is correct if the person should have been counted in the state.

### **8.4.3 Person Omissions**

An overall estimate of the number of people omitted by the census will be obtained by subtracting the total estimated number of correct enumerations from the dual system estimate:

$$Omissions = DSE - CE$$

## **8.5 Missing Data for P-Sample HUs**

There are three missing data operations for HUs in the P sample: imputing HU characteristics, imputing HU status, and imputing HU match status. Each of these will be discussed in more detail in the sections that follow.

### **8.5.1 Imputing Missing Characteristics for P-Sample HUs**

Production of PES HU coverage estimates requires data on occupancy status, tenure (owner versus renter), and race/ethnicity domain of the householder to classify HUs by these important characteristics. So, they will be imputed whenever the data was not collected. The census edit and imputation system will be used to impute the missing characteristics. In general, if the HU is occupied and interviewed in the PES PI, the values used for HU estimation will be the same as those used for person estimation. If the HU was not interviewed in the PES PI, we will assume it is occupied.

### **8.5.2 Imputing Missing HU Status for P-Sample HUs**

The PES HU matching operations must determine if a P-sample HU was a HU on Census Day. For some units, this cannot be determined, and the HU status will be unresolved. We will assign a probability of being a HU status for all sample cases. Cases determined to be a HU will receive a probability of 1. Cases determined not to be HUs will receive a probability of 0. The cases with unresolved status will have a probability assigned to them.

---

<sup>29</sup> The PES coverage universe includes HUs and persons in HUs. It does not include GQs, or remote Alaska enumeration areas.



We fit one or more logistic regression models to the resolved cases. Using these models and the characteristics of the unresolved cases, we can then impute a probability of being a HU for these unresolved cases.

### **8.5.3 Imputing Missing Match Status for P-Sample HUs**

The PES HU Matching operation must determine if the HU matched a census enumeration in the net error search area. This process is simpler for HUs than for people, because mover status is not a consideration. We will assign a probability of match for all sample cases. Matches will receive a probability 1. Nonmatches will receive a probability of 0. The cases with unresolved match status will have a match probability assigned to them.

We will fit one or more logistic regression model to the resolved cases. Using these models and the characteristics of the unresolved cases, we can then impute a probability of match for the unresolved cases.

## **8.6 Missing Data for E-Sample HUs**

There are three missing data procedures needed for HUs in the E sample: a procedure for dealing with missing HU characteristics, a procedure for imputing enumeration status (correct or erroneous) for DSE, and a procedure for estimating components of coverage. Each of these processes will be discussed in more detail in the following sections.

### **8.6.1 Missing Characteristics for E-Sample HUs**

Production of PES HU coverage estimates requires data on occupancy status, tenure (owner versus renter), and race/ethnicity domain of the householder to classify HUs by these important characteristics. So, they will be imputed whenever the data was not collected. For cases in the E sample with missing HU characteristics, we will not do any special imputation procedures. Instead, analogous to the procedure for missing data for E-sample people, we will use the CEF and obtain the imputed characteristics from there.

### **8.6.2 Imputing Missing Enumeration Status for E-Sample HUs (Net Error)**

The probability of being a correct enumeration for net coverage will be determined using logistic regression. Cases with insufficient information for matching and followup will not be used in the modeling.

### **8.6.3 Imputing Missing Enumeration Status for E-Sample HUs (Components)**

When a HU's enumeration status is unresolved, it is unresolved for both net coverage and components of coverage. In this way, the HU estimation is unlike the person estimation, where person enumerations can be resolved for net error, but unresolved for components of error. The imputation for components of coverage is conducted separately from the imputation for net error. There are five components of coverage.

- Correctly enumerated in the sample BCU.
- Correctly enumerated in the surrounding ring of BCUs.
- Correctly enumerated outside the ring of surrounding BCUs.
- Erroneously enumerated due to duplication.
- Erroneously enumerated due to reasons other than duplication.

To satisfy the assumptions of the DSE and to properly estimate net coverage, HUs correctly enumerated outside the surrounding ring of BCUs' are classified as erroneous enumerations for net coverage, even though they are classified as correct enumerations for components of coverage.

### 8.7 HU Net Coverage Error

There are no high-quality data sources that can be used for a correlation bias adjustment of housing units. Thus, the dual system estimate for HUs does not include the correlation bias adjustment. Furthermore, the data-defined concept does not apply to HUs, so the associated predicted probability of being data-defined is also excluded from the DSE formula. The DSE for HUs can be expressed as

$$DSE = \sum_{j \in C} \frac{\pi_{ce(j)}}{\pi_{m(j)}}$$

where  $j$  indexes housing units enumerated in the Census. The predicted correct enumeration and match probabilities ( $\pi_{ce}$ ,  $\pi_m$ , respectively) will be obtained through logistic regression modeling.

We will then estimate net coverage by comparing the estimate of the true population (from the dual system estimator) to the census HU count.

### 8.8 HU Components of Census Coverage

The process for computing component estimates for HUs is very similar to the one used for people, including a ratio adjustment as described in Section 9.4.1. The ratio adjustment cells can be based on region, occupancy status or other variables and the cell totals are the corresponding census HU counts. Depending on sample sizes and the size of the adjustment factors, some collapsing of the original cells may be necessary.

For HUs, separate estimates of correct enumerations, erroneous enumerations, and omissions will be computed. The correct enumerations will be of two kinds: correctly enumerated in the BCU and correctly enumerated in the surrounding BCUs. The erroneous enumeration estimate will be broken down into three parts: structures enumerated in the census as HUs, but did not exist or were not HUs; HUs enumerated more than once (duplicates); and geocoding errors.



As with person omissions, the estimated number of HUs omitted from the census will be obtained by subtracting the estimate of correct enumerations from the DSE:

$$Omissions = DSE - CE$$

## 8.9 Uncertainty Estimation

For all estimates of correct and erroneous enumerations, we will publish standard errors based on successive difference replicates using 80 replicates assigned during the BCU sample selection.

For national-level estimates of net coverage, we will repeat all estimation operations using survey weights for each successive difference replicate. For example, we will fit a survey-weighted logistic regression model for the match rate within each replicate, and use each of the 80 models to predict 80 match rates for each census record.

For biased synthetic estimates of net coverage, we will calculate mean squared errors. As we did for the 2010 CCM, we plan to use the smoothed group method to estimate the mean squared error of biased synthetic estimates. However, we do not plan to disseminate sub-state estimates. And, we plan to produce unbiased state estimates of net coverage for states. Thus, we do not have plans to disseminate any biased synthetic estimates.

## 9. References

- Anderson, F. and Hartman, P. (2009). "2010 Census Coverage Measurement Initial Housing Unit and FHU Clerical Matching Quality Assurance Plan – Revision." . DSSD2010 Census Coverage Measurement Memorandum Series #2010-D5-02R.
- Beaghen, M. (Forthcoming). "High-Level Design Requirements for 2020 Post-Enumeration Survey Final Housing Unit Computer Processing and Matching, Clerical Matching, and Followup." DSSD 2020 Post-Enumeration Survey Memorandum Series #2020-B-07.
- Cantwell, Patrick (2018). "Independence Rules for the 2020 Post-Enumeration Survey and 2020 Census Operations Control Systems for Field Data Collection." DSSD 2020 Post-Enumeration Survey Memorandum Series, Draft August 24, 2017.
- Cecchi, R. (2008). "Quality Assurance Plan for the 2010 Census Coverage Measurement Independent Listing Field Operation." DSSD 2010 Census Coverage Measurement Memorandum Series #2010-D5-01.
- Cecchi, R. (2009). "Quality Assurance Plan for the 2010 Census Coverage Measurement Initial HU Followup Operation – Revision." DSSD 2010 Census Coverage Measurement Memorandum Series #2010-D5-06R.
- Cecchi, R. (2010a). "Quality Assurance Plan for the 2010 Census Coverage Measurement Person Followup Field Operation." DSSD 2010 Census Coverage Measurement Memorandum Series #2010-D10-14.
- Cecchi, R. (2010b). "Quality Assurance Plan for the 2010 Census Coverage Measurement FHU Followup Operation." DSSD 2010 Census Coverage Measurement Memorandum Series #2010-D5-13.
- Cronkite, D. and Shoemaker, H. (2010). "Codes Used In 2010 Census Coverage Measurement Initial HU Matching." DSSD Census Coverage Measurement Memorandum Series #2010-B-12R1.
- Davis, P. (2008). "2010 Census Coverage Measurement: Block Cluster Sample for Independent Listing Software Requirements Specification." DSSD 2010 Census Coverage Measurement Memorandum Series #2010-C-05R2.
- Davis, P. (2005). "2010 Census Coverage Measurement: Provisional Sample Allocation." DSSD 2010 Census Coverage Measurement Memorandum Series #2010-C-01.
- Fay, R.E., Passel, J.S., and Robinson, J.G. with assistance from C.D. Cowan (1988). *The Coverage of Population in the 1980 Census*, 1980 Census of Population and Housing, Evaluation and Research Reports PHC80-E4. Washington, D.C.: U.S. Department of Commerce.



- Hartman, P. (2010). "Quality Assurance Plan for the 2010 Census Coverage Measurement Person Interview Field Operation – Revision." DSSD 2010 Census Coverage Measurement Memorandum Series #2010-D10-07R.
- Hill, C., Nguyen, T., Davis, L. (2018). "2020 Post-Enumeration Survey: Sample Design," DSSD 2020 POST-ENUMERATION SURVEY MEMORANDUM SERIES #2020-C-08.
- Hogan, H. (1993). "The 1990 Post-Enumeration Survey: Operations and Results." *Journal of the American Statistical Association*, Vol. 88, No. 423, 1047-1060.
- Johnson, S. and Adams, T. (2009). "2010 Census Coverage Measurement: Codes Used in Person Matching." DSSD 2010 Census Coverage Measurement Memorandum Series #2010-B-11.
- Kjeldgaard, I. (Forthcoming). "High-Level Design Requirements for 2020 Post-Enumeration Survey Independent Listing Operation." DSSD 2020 Post-Enumeration Survey Memorandum Series #2020-B-03.
- Kraetsch, M. (Forthcoming). "High-Level Design Requirements for 2020 Post-Enumeration Survey Initial Housing Unit Matching and Followup Operations." DSSD 2020 Post-Enumeration Survey Memorandum Series #2020-B-04.
- Kraetsch, M. (Forthcoming). "High-Level Design Requirements for 2020 Post-Enumeration Survey Person Interview and Reinterview." DSSD 2020 Post-Enumeration Survey Memorandum Series #2020-B-05.
- Konicki, S. (2010). "2010 Census Coverage Measurement: Sample Design – Revised." DSSD 2010 Census Coverage Measurement Memorandum Series #2010-C-13R1.
- Mule, T. (2002). "A.C.E. Revision II Results: Further Study of Person Duplication," DSSD Revised A.C.E. Memorandum Series PP-51.
- Mule, T. (2008). "2010 Census Coverage Measurement Estimation Methodology." DSSD 2010 Census Coverage Measurement Memorandum Series #2010-E-18.
- Mulligan, J. (2009). "2010 Census Coverage Measurement: 2010 Block Cluster Reduction Results." DSSD 2010 Census Coverage Measurement Memorandum Series #2010-C-17.
- Nguyen, T. (Forthcoming). "High-Level Design Requirements for 2020 Post-Enumeration Survey Sampling Operations." DSSD 2020 Post-Enumeration Survey Memorandum Series #2020-B-02.
- Shoemaker, H. (2008a). "High-Level Design Requirements for 2010 Census Coverage Measurement Independent Listing Operation." DSSD 2010 Census Coverage Measurement Memorandum Series #2010-B-04.

- Shoemaker, H. (2008b). "High-Level Requirements for the 2010 Census Coverage Measurement Person Interview and Reinterview." DSSD 2010 Census Coverage Measurement Memorandum Series #2010-B-06.
- Shoemaker, H. (2008c). "High-Level Design Requirements for 2010 Census Coverage Measurement Person Matching and Followup." DSSD 2010 Census Coverage Measurement Memorandum Series #2010-B-09R1.
- Shoemaker, H. (2008d). "High-Level Design Requirements for 2010 Census Coverage Measurement FHU Processing, Matching, and Followup." DSSD 2010 Census Coverage Measurement Memorandum Series #2010-B-08.
- Shoemaker, H. (2009). "2010 Census Coverage Measurement Glossary." DSSD 2010 Census Coverage Measurement Memorandum Series #2010-B-10.
- Shoemaker, H. (2011). "High-Level Design Requirements for 2010 Census Coverage Measurement Initial HU Matching and Followup-REVISION#2." DSSD 2010 Census Coverage Measurement Memorandum Series #2010-B-05R2.
- Singh, R. (2005). "2010 Census Coverage Measurement – Updated Plans." DSSD 2010 Census Coverage Measurement Memorandum Series #A-06.
- Thompson, A. (Forthcoming). "High-Level Design Requirements for 2020 Post-Enumeration Survey Person Matching and Followup Operations." DSSD 2020 Post-Enumeration Survey Memorandum Series #2020-B-06.
- U.S. Census Bureau (2004). "Accuracy and Coverage Evaluation of Census 2000: Design and Methodology."
- Vitrano, F. (2009). "Residence Rule and Residence Situations for the 2010 Census." 2010 Census Information Memorandum Series No. 35, June 1, 2009.
- Whitford, D. (2008). "2010 Census Coverage Measurement: Block Cluster Delineation." DSSD 2010 Census Coverage Measurement Memorandum Series #2010-C-03-R1.
- Whitford, D. (2009). "Recommendation to Reduce Nonsampling Error in the 2010 Census Coverage Measurement Program." DSSD 2010 Census Coverage Measurement Memorandum Series #A-43.
- Zamora, J. (Forthcoming). "High-Level Design Requirements for 2020 Post-Enumeration Survey Estimation Operation." DSSD 2020 Post-Enumeration Survey Memorandum Series #2020-B-08.

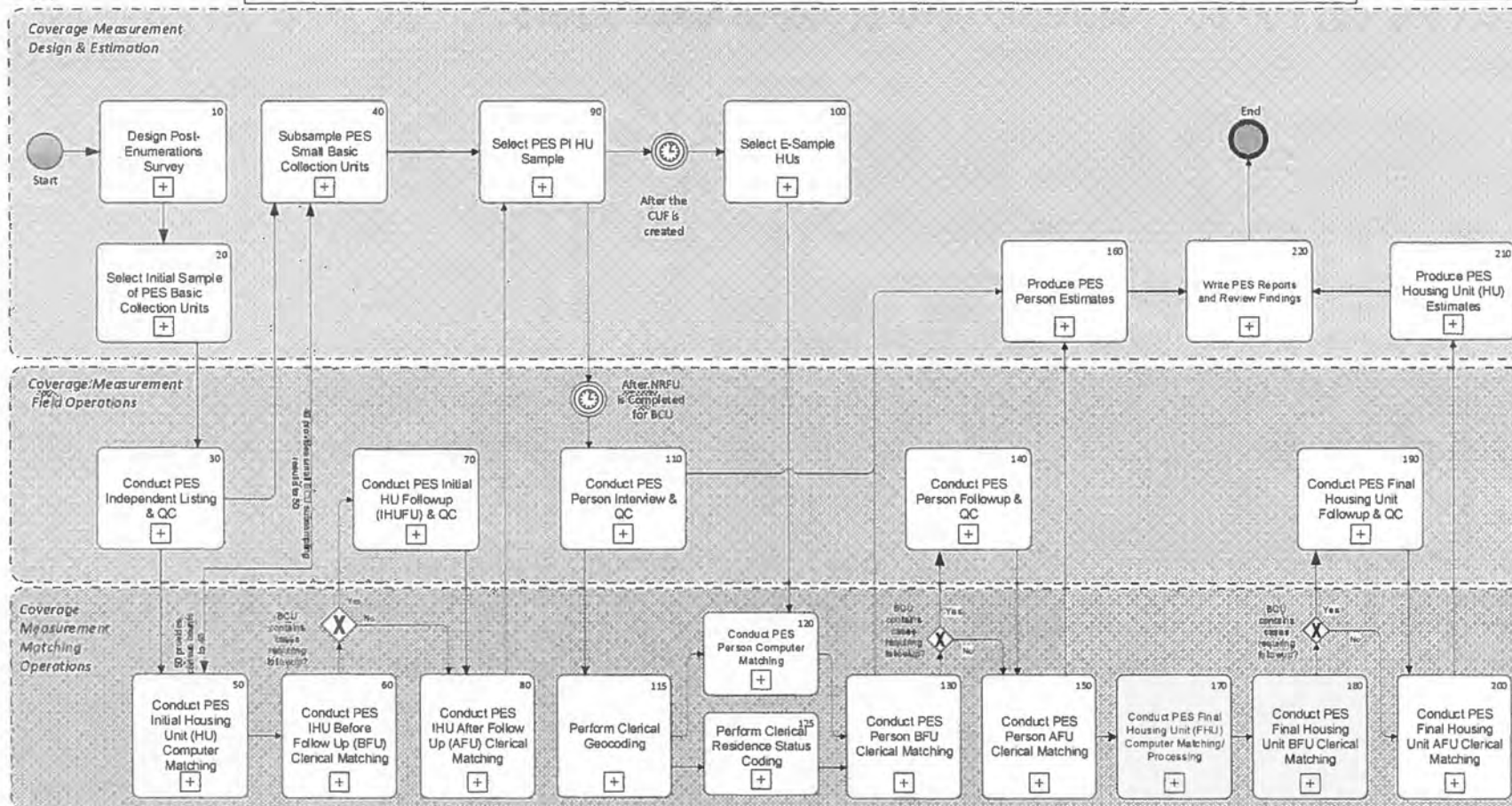






## 2020 Coverage Measurement (CM)

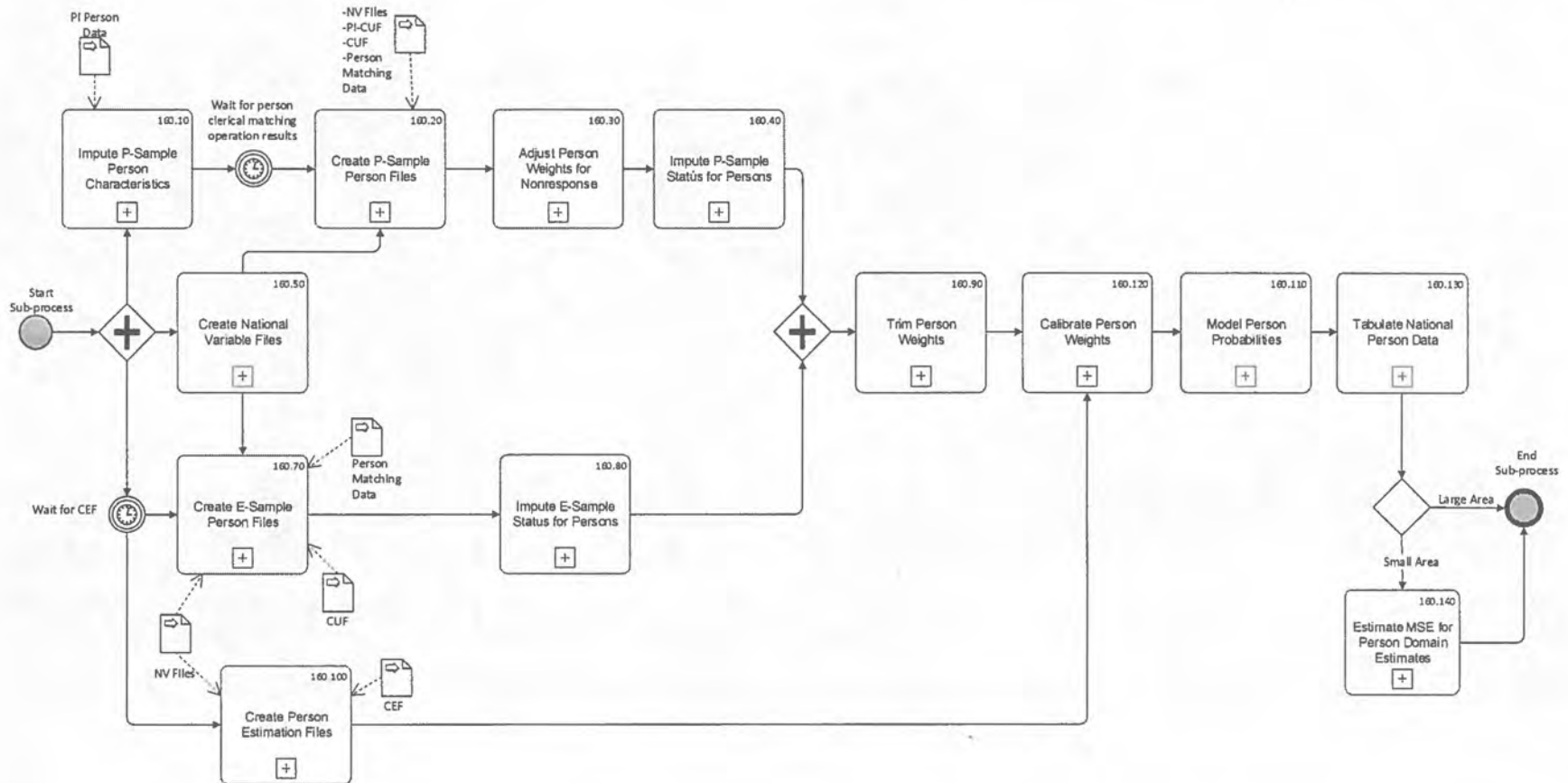
Purpose: To evaluate coverage of the 2020 Census in order to improve future censuses. The CM Program conducts the Post-Enumeration Survey (PES) to measure the coverage of the 2020 Census. The PES is designed to measure the coverage of housing units and people, excluding group quarters and people residing in group quarters.





## 160. Produce PES Person Estimates

Purpose: The PES person estimation produces tables and estimates of person coverage for the decennial census.



## 210. Produce PES Housing Unit Estimates

Purpose: The PES Housing Unit estimation produces tables and estimates of housing unit coverage for the decennial census.

